

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2001, 16:50:26 ; Search time 22.51 Seconds
(without alignments)
16.920 Million cell updates/sec

Title: US-09-251-638-2

Perfect score: 27

Sequence: 1 GVGVP 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	38	T46593	phytoene dehydroge
2	27	100.0	84	S03384	hypothetical prote
3	27	100.0	97	C96743	unknown protein (i
4	27	100.0	161	S33491	hypothetical prote
5	27	100.0	206	T10113	hypothetical prote
6	27	100.0	208	D64301	hypothetical prote
7	27	100.0	211	S73791	hypothetical prote
8	27	100.0	214	T23593	hypothetical prote
9	27	100.0	256	G39845	hypothetical prote
10	27	100.0	277	S78063	homeobox protein p
11	27	100.0	281	JC4558	dihydroorotate deh
12	27	100.0	290	S67297	viomycin kinase (E
13	27	100.0	303	F70693	hypothetical prote
14	27	100.0	308	B5292	probable ugpa prot
15	27	100.0	336	T21461	glucokinase - Dein
16	27	100.0	346	A64475	hypothetical prote
17	27	100.0	367	C84236	hypothetical prote
18	27	100.0	372	B48227	delta opioid recep
19	27	100.0	372	S34592	delta opioid recep
20	27	100.0	372	T38532	delta opioid recep
21	27	100.0	373	T21955	hypothetical prote
22	27	100.0	376	A26066	segmentation prote
23	27	100.0	378	E83373	adenylate cyclase
24	27	100.0	387	S75050	IMP dehydrogenase
25	27	100.0	398	S74939	hypothetical prote
26	27	100.0	399	T18566	hypothetical prote
27	27	100.0	404	E70218	IMP dehydrogenase
28	27	100.0	414	T21954	hypothetical prote
29	27	100.0	422	E83083	probable two-compo

30 27 100.0 426 2 T38650 probable phosphome
31 27 100.0 444 2 T15907 hypothetical prote
32 27 100.0 451 2 D71327 probable D-alanine
33 27 100.0 482 2 C72864 IMP dehydrogenase
34 27 100.0 484 1 B69056 IMP dehydrogenase
35 27 100.0 485 1 G81308 IMP dehydrogenase
36 27 100.0 485 1 JC4998 IMP dehydrogenase
37 27 100.0 485 2 B82558 IMP dehydrogenase
38 27 100.0 485 2 E75015 IMP dehydrogenase
39 27 100.0 486 1 E71456 IMP dehydrogenase
40 27 100.0 487 1 H81109 IMP dehydrogenase
41 27 100.0 487 2 F81906 IMP dehydrogenase
42 27 100.0 488 1 DEECIP IMP dehydrogenase
43 27 100.0 488 1 H64055 IMP dehydrogenase
44 27 100.0 488 1 DEBSMP IMP dehydrogenase
45 27 100.0 488 2 G85894 IMP dehydrogenase

ALIGNMENTS

RESULT 1

T46593
phytoene dehydrogenase [imported] - Mycobacterium marinum (fragment)
C:Species: Mycobacterium marinum
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000
C:Accession: T46593
R:Ramakrishnan, L.; Tran, H.T.; Federspiel, N.A.; Falkow, S.
J. Bacteriol. 179, 5862-5868, 1997
A:Title: A crtB homolog essential for photochromogenicity in Mycobacterium marinum: 1
A:Reference number: Z23096; MUID:97440138
A:Accession: T46593
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-38 <RAM>
A:Cross-references: EMBL:U92075; NID:g1928930; PIDN:RAB71427.1; PID:g1928931
A:Experimental source: strain M
C:Genetics:
A:Gene: crtB

Query Match 100.0%; Score 27; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
|||||
DB 3 GVGVP 7

RESULT 2

S03384
hypothetical protein (IGF-II 3' region) - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 05-Nov-1999
C:Accession: S03384
R:de Pagter-Holthuizen, P.; Jansen, M.; van der Kammen, R.A.; van Schaik, F.M.A.; Sus
Biochim. Biophys. Acta 950, 282-295, 1988
A:Title: Differential expression of the human insulin-like growth factor II gene. Cha
A:Reference number: S03383; MUID:89000779
A:Accession: S03384
A:Molecule type: DNA
A:Residues: 1-84 <DEP>
A:Cross-references: EMBL:X07868; NID:g32998; PIDN:CAA30718.1; PID:g33000
C:Genetics:
A:Map position: 11

Query Match 100.0%; Score 27; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5

Db 76 GVGVP 80
|||||

RESULT 3
C96743
unknown protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96743
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: C96743
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-97 <STO>
A:Cross-references: GB:AE005173; NID:g11054413; PIDN:RAG27800.1; GSPDB:GN00141
C:Genetics:
A:Gene: F28P5.9
A:Map position: 1

Query Match 100.0%; Score 27; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GVGVP 5
|||||
Db 79 GVGVP 83

RESULT 4
S33491
hypothetical protein 6 (tytel) - fowl adenovirus 1
C:Species: Aviadenovirus gall (fowl adenovirus 1, CELO)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
C:Accession: S33491
R:Akopian, T.A.; Kaverina, E.N.; Kruglyak, V.A.; Naroditsky, B.S.; Tikhonenko, T.T.
submitted to the EMBL Data Library, May 1993
A:Description: Sequence of an avian adenovirus (CELO) DNA fragment (11.2 - 19.2 %).
A:Reference number: S33486
A:Accession: S33491
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-161 <AKO>
A:Cross-references: EMBL:Z22864; NID:g311514; PIDN:CAA80482.1; PID:g311520
C:Superfamily: fowl adenovirus 1 hypothetical protein 6

Query Match 100.0%; Score 27; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GVGVP 5
|||||
Db 57 GVGVP 61

RESULT 5
T10113
hypothetical protein - maize streak virus (isolate SP2)
C:Species: maize streak virus
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10113

R:Isnard, M.; Granier, M.; Frutos, R.; Reynaud, B.; Peterschmitt, M.
J. Gen. Virol. 79, 3091-3099, 1998
A:Title: Quasispecies nature of three related maize streak virus isolates obtained th
A:Reference number: Z16955; MUID:99094636
A:Accession: T10113
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-206 <ISN>
A:Cross-references: EMBL:AJ225008; NID:g2980718; PIDN:CAA12317.1; PID:g2980722

Query Match 100.0%; Score 27; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GVGVP 5
|||||
Db 128 GVGVP 132

RESULT 6
D64301
hypothetical protein MJ0012 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 10-Oct-1997
C:Accession: D64301
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A:Reference number: A64300; MUID:96337999
A:Accession: D64301
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-208 <BUL>
A:Cross-references: GB:U67460; GB:L77117; NID:g1590820; PID:g1590823; TIGR:WJ0012; PI
C:Genetics:
A:Map position: REV14050-13424
A:Start codon: TTG

Query Match 100.0%; Score 27; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GVGVP 5
|||||
Db 130 GVGVP 134

RESULT 7
S73791
hypothetical protein A19_ORF211 - Mycoplasma pneumoniae (strain ATCC 29342)
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
C:Accession: S73791
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirk1, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumon
A:Reference number: S73327; MUID:97105885
A:Accession: S73791
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-211 <HIM>
A:Cross-references: EMBL:AE000046; GB:U00089; NID:g1674152; PIDN:AAB96113.1; PID:g167
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: Mycoplasma pneumoniae hypothetical protein A19_ORF211

Query Match 100.0%; Score 27; DB 2; Length 211;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 Db 24 GVGVP 28
 |||||

RESULT 8
 T23593
 hypothetical protein K10H10.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000
 C:Accession: T23593
 R:Percy, C.
 submitted to the EMBL Data Library, December 1996
 A:Reference number: Z19766
 A:Accession: T23593
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-214 <WIL>
 A:Cross-references: EMBL:283236; PIDN:CAB05780.1; GSPDB:GN00020; CBSP:K10H10.4
 A:Experimental source: clone K10H10
 C:Genetics:
 A:Gene: CBSP:K10H10.4
 A:Map position: 2
 A:Introns: 117/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein K10H10.4

Query Match 100.0%; Score 27; DB 2; Length 214;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 Db 195 GVGVP 199
 |||||

RESULT 9
 G39845
 dihydroorotate dehydrogenase (electron transfer subunit) pyrDII - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 20-Mar-1992 #sequence_revision 10-Apr-1992 #text_change 21-Jul-2000
 C:Accession: G39845; D69686
 R:Quinn, C.L.; Stephenson, B.T.; Switzer, R.L.
 J. Biol. Chem. 266, 9113-9127, 1991
 A:Title: Functional organization and nucleotide sequence of the Bacillus subtilis pyrDII
 A:Reference number: A39845; MUID:91225016
 A:Accession: G39845
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-256 <QUI>
 A:Cross-references: GB:M59757; NID:94887706; PIDN:AAA1271.1; PID:G143391
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertain
 C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caidwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
 Koester, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033
 A:Accession: D69686

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-256 <KUN>
 A:Cross-references: GB:299112; GB:AL009126; NID:92633902; PIDN:CABL3427.1; PID:9263339
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: pyrDII
 C:Superfamily: Pyrococcus furiosus cytochrome-c3 hydrogenase gamma chain

Query Match 100.0%; Score 27; DB 2; Length 256;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 Db 116 GVGVP 120
 |||||

RESULT 10
 S78063
 homeobox protein prh - chicken
 C:Species: Gallus gallus (chicken)
 C>Date: 19-Nov-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1999
 C:Accession: S78063
 R:Crompton, M.R.; Bartlett, T.J.; MacGregor, A.D.; Manfioletti, G.; Buratti, E.; Gian
 Nucleic Acids Res. 20, 5661-5667, 1992
 A:Title: Identification of a novel vertebrate homeobox gene expressed in haematopoiet
 A:Reference number: S26799; MUID:93087175
 A:Accession: S78063
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-277 <CRO>
 A:Cross-references: EMBL:X64711; NID:9297086; PIDN:CAA45966.1; PID:9297087
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:145-201/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 27; DB 2; Length 277;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 Db 15 GVGVP 19
 |||||

RESULT 11
 JC4558
 viomycin kinase (EC 2.7.1.103) - Streptomyces capreolus
 N:Alternate names: capreomycin phosphotransferase; CPH protein; viomycin phosphotrans
 C:Species: Streptomyces capreolus
 C>Date: 12-Mar-1996 #sequence_revision 19-Apr-1996 #text_change 05-Nov-1999
 C:Accession: JC4558
 R:Thiara, A.S.; Cundliffe, E.
 Gene 167, 121-126, 1995
 A:Title: Analysis of two capreomycin-resistance determinants from Streptomyces capreo
 A:Reference number: JC4557; MUID:96144260
 A:Accession: JC4558
 A:Molecule type: DNA
 A:Residues: 1-281 <THI>
 A:Cross-references: GB:U13078; NID:95331118; PIDN:AAA92037.1; PID:9533119
 A:Experimental source: NCIMB 9801
 C:Comment: This enzyme is a capreomycin-modifying enzyme, and it inactivates the comp
 B and IIB.
 C:Genetics:
 A:Gene: cph
 C:Keywords: phosphotransferase

Query Match 100.0%; Score 27; DB 2; Length 281;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5

Db 68 GVGVP 72

RESULT 12

S67297

hypothetical protein YOR385w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein O6768

C:Species: Saccharomyces cerevisiae

C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 15-Sep-2000

A:Accession: S67297

R:Delius, H.; Hebling, U.; Hofmann, B.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67261

A:Accession: S67297

A:Molecule type: DNA

A:Residues: 1-290

A:Cross-references: EMBL:275293; NID:g1420828; PID:e252450; GSPDB:GN00015;

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YOR385w

A:Map position: 15R

C:Superfamily: Saccharomyces cerevisiae hypothetical protein YMR316w

Query Match

Best Local Similarity 100.0%; Score 27; DB 2; Length 290;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5

Db 74 GVGVP 78

RESULT 13

F70693

probable upgA protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

A:Accession: F70693

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987

A:Accession: F70693

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-303 <COL>

A:Cross-references: GB:Z81331; GB:AL123456; NID:g3261650; PID:CA803651.1; PID:g1648885

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: upgA

C:Superfamily: inner membrane protein upgA

Query Match

Best Local Similarity 100.0%; Score 27; DB 2; Length 303;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5

Db 153 GVGVP 157

RESULT 14

B75292

glucokinase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jun-2000

A:Accession: B75292

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: B75292

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <WHI>

A:Cross-references: GB:AE002061; GB:AE000513; NID:g6460095; PIDN:AAFI1841.1; PID:g646

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2296

A:Map position: 1

C:Superfamily: glucose kinase; glucose kinase homology

F:67-201/Domain: glucose kinase homology <GRK>

Query Match 100.0%; Score 27; DB 2; Length 308;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5

Db 71 GVGVP 75

RESULT 15

T21461

hypothetical protein F27E5.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000

A:Accession: T21461; T21709

R:Wilkinson, J.

submitted to the EMBL Data Library, March 1995

A:Reference number: Z19425

A:Accession: T21461

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-336 <WIL>

A:Cross-references: EMBL:Z48582; PIDN:CAA88468.1; GSPDB:GN00020; CESP:F27E5.2

A:Experimental source: clone F27E5

R:Wilkinson, J.

submitted to the EMBL Data Library, March 1995

A:Reference number: Z19462

A:Accession: T21709

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-336 <WI2>

A:Cross-references: EMBL:Z48783; PIDN:CAA88702.1; GSPDB:GN00020; CESP:F27E5.2

A:Experimental source: clone F33H1

C:Genetics:

A:Gene: CESP:F27E5.2

A:Map position: 2

A:Introns: 14/3; 47/2; 79/3; 141/1; 166/1; 264/1

C:Superfamily: paired box transcription factor Pax-4; homeobox homology; paired box h

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:13-138/Domain: paired box homology <PBH>

Query Match 100.0%; Score 27; DB 2; Length 336;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5

Db 141 GVGVP 145

Search completed: July 25, 2001, 16:50:27
Job time: 98 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2001, 16:51:25 ; Search time 12.69 seconds
(without alignments)
13.497 Million cell updates/sec

Title: US-09-251-638-2

Perfect score: 27

Sequence: 1 GVGVP 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	27	100.0	84	IG2R_HUMAN
2	27	100.0	208	Y012_METJA
3	27	100.0	211	YD71_MYCPN
4	27	100.0	228	OPRD_PIG
5	27	100.0	256	PRK_BACSU
6	27	100.0	277	HPKH_CHICK
7	27	100.0	291	Y4TQ_RHISN
8	27	100.0	372	OPRD_HUMAN
9	27	100.0	372	OPRD_MOUSE
10	27	100.0	372	OPRD_RAT
11	27	100.0	376	HMEV_DROME
12	27	100.0	404	IMDH_BORBU
13	27	100.0	477	ME12_HUMAN
14	27	100.0	477	ME12_MOUSE
15	27	100.0	485	IMDH_PYRPU
16	27	100.0	486	IMDH_PYRHO
17	27	100.0	488	IMDH_ECOLI
18	27	100.0	488	IMDH_HAETIN
19	27	100.0	490	IMDH_AQUAE
20	27	100.0	492	IMDH_STRPY
21	27	100.0	496	IMDH_METJA
22	27	100.0	507	CRTI_STRGR
23	27	100.0	507	DAF_CAVPO
24	27	100.0	508	CRTI_STRSE
25	27	100.0	513	IMDH_BACSU
26	27	100.0	635	CA28_HUMAN
27	27	100.0	684	FLID_HELPY
28	27	100.0	730	ELIS_HUMAN
29	27	100.0	747	ELIS_BOVIN
30	27	100.0	747	YNS4_CAEEL
31	27	100.0	750	ELIS_CHICK
32	27	100.0	775	YTX1_XENLA
33	27	100.0	1758	CA24_CAEEL

34	27	100.0	1805	1	RW1_HUMAN	Q92545 homo sapien
35	27	100.0	1829	1	RW1_MOUSE	Q70472 mus musculu
36	27	100.0	2944	1	CA17_HUMAN	Q02388 homo sapien
37	26	96.3	65	1	Y12J_BPT4	Q02406 bacterioph
38	26	96.3	125	1	KCIA_PIG	Q19175 sus scrofa
39	26	96.3	139	1	HMAA_ARTSF	Q05007 artemia san
40	26	96.3	146	1	SP22_BACLI	P26778 bacillus li
41	26	96.3	146	1	SP22_BACST	Q32727 bacillus st
42	26	96.3	198	1	LEUD_MYCLE	Q33134 mycobacteri
43	26	96.3	198	1	LEUD_MYCTU	O53236 mycobacteri
44	26	96.3	200	1	TFAB_ECOLI	P09133 escherichia
45	26	96.3	203	1	TFAB_ECOLI	Q47427 escherichia

ALIGNMENTS

RESULT	1
IG2R_HUMAN	
ID	IG2R_HUMAN STANDARD; PRT; 84 AA.
AC	P09565;
DT	01-MAR-1989 (Rel. 10, Created)
DT	01-MAR-1989 (Rel. 10, Last sequence update)
DT	01-MAR-1989 (Rel. 10, Last annotation update)
DE	PUTATIVE INSULIN-LIKE GROWTH FACTOR II ASSOCIATED PROTEIN.
OS	Homo sapiens (Human)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID:9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Placenta;
RX	MEDLINE=89000779; PubMed=3167054;
RA	de Pagter-Holthuisen P., van der Kammen R.A., Jansen M.,
RA	van Schaik F.M.A., Sussenbach J.S.;
RT	"Differential expression of the human insulin-like growth factor II
RT	gene. Characterization of the IGF-II mRNAs and an mRNA encoding a
RT	putative IGF-II-associated protein";
RL	Biochim. Biophys. Acta 950:282-295(1988).
CC	-1- FUNCTION: NOT KNOWN.
CC	-1- MISCELLANEOUS: THE COORDINATED EXPRESSION OF INSULIN-LIKE GROWTH
CC	FACTOR (IGF II) ASSOCIATED PROTEIN AND IGF II IS DEVELOPMENTALLY
CC	AND TISSUE-SPECIFICALLY REGULATED.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; X07868; CAA30718.1; -
DR	PIR; S03384; S03384.
KW	Growth factor
SQ	SEQUENCE 84 AA; 9081 MW; 8BFBB4E49E5EFAB CRC64;

Query Match 100.0%; Score 27; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVGVP 5
|||||
Db 76 GVGVP 80

RESULT	2
Y012_METJA	
ID	Y012_METJA STANDARD; PRT; 208 AA.
AC	Q60328;
DT	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HYPOTHETICAL PROTEIN MJ0012.
GN MJ0012.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glöck A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG, TO THE C-TERMINAL OF M.JANNASCHII MJ1635.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D67460; -; NOT_ANNOTATED_CDS.
DR TIGR; MJ0012; -;
KW Hypothetical protein.
SQ SEQUENCE 208 AA; 24347 MW; 4C824F93015561B4 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVGVP 5
Db 130 GVGVP 134
|||||

RESULT 3
YD71_MYCPN STANDARD; PRT; 211 AA.
AC P75410;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN MPN371 (A19_ORF211).
GN MPN371 OR MP465.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000046; AAB96113.1; -;
KW Hypothetical protein.
SQ SEQUENCE 211 AA; 23592 MW; A5E240288852DDB2 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVGVP 5
Db 24 GVGVP 28
|||||

RESULT 4
OPRD_PIG STANDARD; PRT; 228 AA.
ID OPRD_PIG
AC P79291;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DELTA-TYPE OPIOID RECEPTOR (DOR-1) (FRAGMENT).
GN OPRD1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cortex;
RX MEDLINE=98352682; PubMed=9690372;
RA Brown D.R., Poonyachoti S., Osinski M.A., Kowalski T.R.,
RA Pampusch M.S., Elde R.P., Murtaugh M.P.;
RT "Delta-opioid receptor mRNA expression and immunohistochemical
localization in porcine ileum.";
RL Dig. Dis. Sci. 43:1402-1410(1998).
CC -1- FUNCTION: INHIBITS NEUROTRANSMITTER RELEASE BY REDUCING CALCIUM
CC ION CURRENTS AND INCREASING POTASSIUM ION CONDUCTANCE. HIGHLY
CC STEREOSELECTIVE RECEPTOR FOR ENKEPHALINS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U71149; AAB39694.1; -;
CC GCRDB; GCR_1526; -;
DR InterPro; IPR000276; -;
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_FL_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEPTOR_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT NON_TER 1 1
FT TRANSMEM <1 3 1 (POTENTIAL).
FT DOMAIN 4 12 2 (POTENTIAL).
FT TRANSMEM 13 30 2 (POTENTIAL).
FT DOMAIN 31 52 2 (POTENTIAL).
FT TRANSMEM 53 72 3 (POTENTIAL).
FT DOMAIN 73 102 4 (POTENTIAL).
FT TRANSMEM 103 118 4 (POTENTIAL).
FT DOMAIN 119 143 5 (POTENTIAL).
FT TRANSMEM 144 166 5 (POTENTIAL).

FT DOMAIN 167 189 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 190 212 6 (POTENTIAL).
 FT DOMAIN 213 221 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 222 >228 7 (POTENTIAL).
 FT DISULFID 49 126 BY SIMILARITY.
 FT NON_TER 228 228
 SQ SEQUENCE 228 AA; 25725 MW; C8C1A984AA711DE CRC64;

Query Match 100.0%; Score 27; DB 1; Length 228;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 |||||
 DB 106 GVGVP 110

RESULT 5
 PYRK_BACSU
 ID PYRK_BACSU STANDARD; PRT; 256 AA.
 AC P25983;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DIHYDROOROTATE DEHYDROGENASE ELECTRON TRANSFER SUBUNIT.
 GN PYRK OR PYR2 OR PYRDB OR PYRDI.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91225016; PubMed=1709162;
 RA Quinn C.L., Stephenson B.T., Switzer R.L.;
 RT "Functional organization and nucleotide sequence of the Bacillus
 RT subtilis pyrimidine biosynthetic operon.";
 RL J. Biol. Chem. 266:9113-9127(1991).
 RN [2]
 RP FUNCTION.
 RC STRAIN=168 / DB104;
 RX MEDLINE=96326349; PubMed=8759868;
 RA Kahler A.E., Switzer R.L.;
 RT "Identification of a novel gene of pyrimidine nucleotide
 RT biosynthesis, pyrDII, that is required for dihydroorotate
 RT dehydrogenase activity in Bacillus subtilis.";
 RL J. Bacteriol. 178:5013-5016(1996).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=20014515; PubMed=10545205;
 RA Kahler A.E., Nielsen F.S., Switzer R.L.;
 RT "Biochemical characterization of the heteromeric Bacillus subtilis
 RT dihydroorotate dehydrogenase and its isolated subunits.";
 RL Arch. Biochem. Biophys. 371:191-201(1999).
 CC -1- FUNCTION: PROBABLE ELECTRON CARRIER PROTEIN INVOLVED IN THE
 CC TRANSFER OF REDUCING EQUIVALENTS FROM THE FLAVOPROTEIN SUBUNIT
 CC (PYRD) TO THE ELECTRON TRANSPORT SYSTEM IN THE CELL MEMBRANE.
 CC -1- COFACTOR: BINDS A 2FE-2S CLUSTER AND FAD.
 CC -1- PATHWAY: FOURTH STEP IN PYRIMIDINE BIOSYNTHESIS.
 CC -1- SUBUNIT: HETEROTETRAMER OF 2 PYRK AND 2 PYRD SUBUNITS.
 CC -1- SIMILARITY: BELONGS TO THE PYRK FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M59757; AAA1271.1;
 CC EMBL; Z99112; CAB13427.1;
 CC PIR; G39845; G39845.

DR Subtilist; B610717; pyrK.
 KW Pyrimidine biosynthesis; Electron transport; Iron-sulfur;
 KW Flavoprotein; FAD.
 FT NP_BIND 110 122 FAD (POTENTIAL).
 FT METAL 220 220 IRON-SULFUR (2FE-2S) (PROBABLE).
 FT METAL 225 225 IRON-SULFUR (2FE-2S) (PROBABLE).
 FT METAL 228 228 IRON-SULFUR (2FE-2S) (PROBABLE).
 FT METAL 243 243 IRON-SULFUR (2FE-2S) (PROBABLE).
 SQ SEQUENCE 256 AA; 28099 MW; DC7B7605E39C2E15 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 256;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 |||||
 DB 116 GVGVP 120

RESULT 6
 HMPH_CHICK
 ID HMPH_CHICK STANDARD; PRT; 277 AA.
 AC Q05502;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE HOMEBOX PROTEIN PRH.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93087175; PubMed=1360645;
 RA Crompton M.R., Bartlett T.J., Macgregor A.D., Manfioletti G.,
 RA Buratti E., Giancotti V., Goodwin G.H.;
 RT "Identification of a novel vertebrate homeobox gene expressed in
 RT haematopoietic cells.";
 RL Nucleic Acids Res. 20:5661-5667(1992).
 CC -1- FUNCTION: RECOGNIZES THE DNA SEQUENCE 5'-ATTAA-3'. MAY PLAY A ROLE
 CC IN HEMATOPOIETIC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: IN ALL HEMATOPOIETIC TISSUES EXCEPT
 CC PERIPHERAL BLOOD ERYTHROCYTES AND IN THE LIVER AND LUNG.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X64711; CAA45966.1;
 CC HSSP; P22808; LVND.
 CC TRANSFAC; T02091;
 CC InterPro; IPR001356;
 CC Pfam; PF00046; homeobox; 1.
 CC PRINTS; PR00024; HOMEBOX.
 CC PROSITE; PS00027; HOMEBOX_1; 1.
 CC PROSITE; PS00071; HOMEBOX_2; 1.
 CC Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 FT DOMAIN 1 140 PRO-RICH.
 FT DNA_BIND 144 203 HOMEBOX.
 SQ SEQUENCE 277 AA; 30213 MW; BE744C143FB9F9FC CRC64;

Query Match 100.0%; Score 27; DB 1; Length 277;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GVGVP 5
|||||
Db 15 GVGVP 19

```

RESULT 7
Y4TQ_RHISN STANDARD; PRT; 291 AA.
AC Q53197;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE PROBABLE PEPTIDE ABC TRANSPORTER PERMEASE PROTEIN Y4TQ.
GN Y4TQ.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freilberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96389014; PubMed=8796346;
RA Freilberg C., Perret X., Broughton W.J., Rosenthal A.;
RT "Sequencing the 500-kb GC-rich symbiotic replicon of Rhizobium sp.
RT NGR234 using dye terminators and a thermostable 'sequenase': a
RT beginning.";
RL Genome Res. 6:590-600(1996).
CC -1- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
CC SYSTEM Y4TQORS FOR A PEPTIDE. PROBABLY RESPONSIBLE FOR THE
CC TRANSLLOCATION OF THE SUBSTRATE ACROSS THE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -1- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE OPPBC
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z68203; CAA92399.1; -
DR EMBL; AE000098; AAB91870.1; -
DR InterPro; IPR000515; -
DR Pfam; PF00528; BPD.transp.1;
DR PROSITE; PS00402; BPD_TRANSP_INN_MEMBER; 1.
KW Hypothetical protein; Transport; Amino-acid transport; Transmembrane;
KW Inner membrane; Plasmid.
FT TRANSMEM 28 48
FT TRANSMEM 92 112 POTENTIAL.
FT TRANSMEM 137 157 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 249 269 POTENTIAL.
SQ SEQUENCE 291 AA; 30910 MW; CD82B7EBB91146E5 CRC64;

```

Query Match 100.0%; Score 27; DB 1; Length 291;
 Best Local Similarity 100.0%; Pred. No. 93;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GVGVP 5
 |||||
 Db 230 GVGVP 234

```

RESULT 8
OPRD_HUMAN STANDARD; PRT; 372 AA.
ID OPRD_HUMAN
AC P41143;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DELTA-TYPE OPIOID RECEPTOR (DOR-1).
GN OPRD1 OR OPRD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cortex, and Striatum;
RX MEDLINE=94260835; PubMed=8201839;
RA Knapp R.J., Malatynska E., Fang L., Li X., Babin E., Nguyen M.,
RA Santoro G., Varga E.V., Hruba V.J., Roeske W.R., Yamamura H.I.;
RT "Identification of a human delta opioid receptor: cloning and
RT expression.";
RL Life Sci. 54:PL463-PL469(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95107267; PubMed=7808419;
RA Simonin F., Befort K., Gaveriaux-Ruff C., Matthes H., Nappay V.,
RA Lannes B., Micheletti G., Kieffer B.;
RT "The human delta-opioid receptor: genomic organization, cDNA cloning,
RT functional expression, and distribution in human brain.";
RL Mol. Pharmacol. 46:1015-1021(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INHIBITS NEUROTRANSMITTER RELEASE BY REDUCING CALCIUM
CC ION CURRENTS AND INCREASING POTASSIUM ION CONDUCTANCE. HIGHLY
CC STEREOSELECTIVE. RECEPTOR FOR ENKEPHALINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U07882; AAA18789.1; -
DR EMBL; U10504; AAA83426.1; -
DR EMBL; AL009181; CAA15671.1; -
DR GCRDB; GCR_1017; -
DR GCRDB; GCR_2055; -
DR MIM; 165195; -
DR InterPro; IPR000276; -
DR InterPro; IPR000321; -
DR InterPro; IPR001418; -
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PRINTS; PR00384; OPIOLDR.
DR PRINTS; PS00525; DELTAOPIOIDR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 45 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 46 75 1 (POTENTIAL).
FT DOMAIN 76 84 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 85 102 2 (POTENTIAL).
FT DOMAIN 103 124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 125 144 3 (POTENTIAL).

```

FT DOMAIN 145 174 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 175 190 4 (POTENTIAL).
 FT DOMAIN 191 215 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 216 238 5 (POTENTIAL).
 FT DOMAIN 239 261 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 262 284 6 (POTENTIAL).
 FT DOMAIN 285 293 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 294 310 7 (POTENTIAL).
 FT DOMAIN 311 372 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 121 198 BY SIMILARITY.
 FT LIPID 333 333 PALMITATE (POTENTIAL).
 FT CONFLICT 27 27 C -> F (IN REF. 1).
 FT CONFLICT 40 41 AR -> P (IN REF. 1).
 FT CONFLICT 348 348 A -> P (IN REF. 1).
 FT CONFLICT 370 370 A -> R (IN REF. 1).
 SQ SEQUENCE 372 AA; 40368 MW; 9D483FC39A2BAE4 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 372;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5
 |||||
 Db 178 GVGVP 182

RESULT 9
 OPND_MOUSE
 ID OPND_MOUSE STANDARD; PRT; 372 AA.
 AC P32300.
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DELTA-TYPE OPIOID RECEPTOR (DOR-1) (K56) (MSL-2).
 GN OPRL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93101664; PubMed=1334555;
 RA Kieffer B.L., Belfort K., Gaveriaux-Ruff C., Hirth C.G.;
 RT "The delta-opioid receptor: isolation of a cDNA by expression cloning
 and pharmacological characterization.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:12048-12052(1992).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93110361; PubMed=1335167;
 RA Evans C.J., Keith D.E. Jr., Morrison H., Magendzo K., Edwards R.H.;
 RT "Cloning of a delta opioid receptor by functional expression.";
 RL Science 258:1952-1955(1992).
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93342064; PubMed=83933575;
 RA Yasuda K., Raynor K., Kong H., Breder C.D., Takeda J., Reisine T.,
 Bell G.I.;
 RT "Cloning and functional comparison of kappa and delta opioid
 receptors from mouse brain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6736-6740(1993).
 [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93391482; PubMed=8397421;
 RA Keith D.E. Jr., Anton B., Evans C.J.;
 RT "Characterization and mapping of a delta opioid receptor clone from
 NG108-15 cells.";
 RL Proc. West. Pharmacol. Soc. 36:299-306(1993).
 [5]
 RP SEQUENCE OF 8-372 FROM N.A.

RX MEDLINE=94022364; PubMed=8415697;
 RA Bdega T., Chin H., Kim K., Jung H.H., Kozak C.A., Klee W.A.;
 RT "Regional expression and chromosomal localization of the delta opiate
 receptor gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:9305-9309(1993).
 [6]
 RN 3D-STRUCTURE MODELING.
 RP MEDLINE=97001837; PubMed=8844829;
 RA Alkorta I., Loew G.H.;
 RT "A 3D model of the delta opioid receptor and ligand-receptor
 complexes.";
 RL Protein Eng. 9:573-583(1996).
 CC -|- FUNCTION: INHIBITS NEUROTRANSMITTER RELEASE BY REDUCING CALCIUM
 ION CURRENTS AND INCREASING POTASSIUM ION CONDUCTANCE. HIGHLY
 STEREOSLECTIVE. RECEPTOR FOR ENKEPHALINS.
 CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -|- TISSUE SPECIFICITY: BRAIN, WITH HIGH CONCENTRATIONS IN THE BASAL
 GANGLIA AND LIMBIC REGIONS.
 CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed, usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 EMBL; L06322; AAA37522.1;
 EMBL; L07271; -; NOT_ANNOTATED_CDS.
 EMBL; L11064; AAA37520.1;
 EMBL; S65335; AAA16009.1;
 EMBL; S66181; AAB28546.1;
 PIR; S37807; S37807.
 PIR; B48227; B48227.
 GCRDB; GCR_0229;
 GCRDB; GCR_0493;
 GCRDB; GCR_0634;
 GCRDB; GCR_0842;
 GMD; MGI:97438; Oprd1.
 InterPro; IPR000276;
 InterPro; IPR000321;
 InterPro; IPR001418;
 Pfam; PF00001; 7tm1.1;
 PRINTS; PR00237; GPCRHHODPSN.
 PRINTS; PR00384; OPIOIDR.
 PRINTS; PR00525; DELTAOPIOIDR.
 PROSITE; PS00237; G_PROTEIN_RECEPTOR_FL_1; 1.
 PROSITE; PS0262; G_PROTEIN_RECEPTOR_FL_2; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein;
 Phosphorylation; Lipoprotein; Palmitate.
 DOMAIN 1 45 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 46 75 1 (POTENTIAL).
 DOMAIN 76 84 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 85 102 2 (POTENTIAL).
 DOMAIN 103 124 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 125 144 3 (POTENTIAL).
 DOMAIN 145 174 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 175 190 4 (POTENTIAL).
 DOMAIN 191 215 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 216 238 5 (POTENTIAL).
 DOMAIN 239 261 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 262 284 6 (POTENTIAL).
 DOMAIN 285 293 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 294 310 7 (POTENTIAL).
 DOMAIN 311 372 CYTOPLASMIC (POTENTIAL).
 CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
 DISULFID 121 198 BY SIMILARITY.
 LIPID 333 333 PALMITATE (POTENTIAL).
 SEQUENCE 372 AA; 40561 MW; BC3B3BBD2E52D3F9 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 372;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGVVP 5
 Db 178 GGVVP 182

RESULT 10

OPRD_RAT
 ID OPD_RAT STANDARD; PRT; 372 AA.
 AC P33533;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE DELTA-TYPE OPIOID RECEPTOR (DOR-1) (OPIOID RECEPTOR A).
 GN OPRD1 OR ROR-A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93351652; PubMed=8394245;
 RA Fukuda K., Kato S., Mori K., Nishi M., Takeshima H.;
 RT "Primary structures and expression from cDNAs of rat opioid receptor
 delta- and mu-subtypes.";
 RL FEBS Lett. 327:311-314(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
 RX MEDLINE=94322412; PubMed=7519274;
 RA Aboud M.E., Noel M.A., Farnsworth J.S., Tao Q.;
 RT "Molecular cloning and expression of a delta-opioid receptor from rat
 brain.";
 RL J. Neurosci. Res. 37:714-719(1994).
 CC -1- FUNCTION: INHIBITS NEUROTRANSMITTER RELEASE BY REDUCING CALCIUM
 ION CURRENTS AND INCREASING POTASSIUM ION CONDUCTANCE. HIGHLY
 STEREOSLECTIVE. RECEPTOR FOR ENKEPHALINS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed, usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D16348; BAA03851.1; -;
 DR EMBL; U00475; AAA19939.1; -;
 DR PIR; S34592; S34592.
 DR GCRDB; GCR_0638; -;
 DR GCRDB; GCR_0805; -;
 DR InterPro; IPR000276; -;
 DR InterPro; IPR000321; -;
 DR InterPro; IPR001418; -;
 DR Pfam; PF00001; 7tm.1.1;
 DR PRINTS; PR00237; GPCRHDOPSN.
 DR PRINTS; PR00384; OPIODR.
 DR PRINTS; PR00525; DELTAOPIOIDR.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1;
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1;
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 Phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 45 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 46 75 1 (POTENTIAL).
 FT DOMAIN 76 84 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 85 102 2 (POTENTIAL).
 FT DOMAIN 103 124 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 125 144 3 (POTENTIAL).
 FT DOMAIN 145 174 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 175 190 4 (POTENTIAL).
 FT DOMAIN 191 215 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 216 238 5 (POTENTIAL).
 FT DOMAIN 239 261 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 262 284 6 (POTENTIAL).
 FT DOMAIN 285 293 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 294 310 7 (POTENTIAL).
 FT DOMAIN 311 372 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 18 18 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 33 33 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 121 198 BY SIMILARITY.
 FT LIPID 333 333 PALMITATE (POTENTIAL).
 SQ SEQUENCE 372 AA; 40449 MW; F578BD1F64C61D50 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 372;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGVVP 5
 Db 178 GGVVP 182

RESULT 11

HMEV_DROME
 ID HMEV_DROME STANDARD; PRT; 376 AA.
 AC P06602; P07667; Q9V5E6;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SEGMENTATION PROTEIN EVEN-SKIPPED.
 GN EVE OR CG2328.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta.
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87051744; PubMed=2877745;
 RA McDonald P.M., Ingham P., Struhl G.;
 RT "Isolation, structure, and expression of even-skipped: a second pair-
 rule gene of Drosophila containing a homeo box.";
 RL Cell 47:721-734(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87218536; PubMed=2884106;
 RA Frasch M., Hoey T., Rushlow C., Doyle H., Levine M.;
 RT "Characterization and localization of the even-skipped protein of
 Drosophila.";
 RL EMBO J. 6:749-759(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Ashburner M., Henderson S.N.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Chen L.X.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazef R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattle B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulo G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klanos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RN Science 287:2185-2195(2000).
RN [4]
RP SEQUENCE OF 1-58 FROM N.A.
RC STRAIN-OR-RC, WA-F, ZM56, AF-S, AND FL-S;
RX MEDLINE=96038621; PubMed=8524036;
RA Ludwig M.Z., Kreitman M.;
RT "Evolutionary dynamics of the enhancer region of even-skipped in
Drosophila.";
RL Mol. Biol. Evol. 12:1002-1011(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 70-189.
RX MEDLINE=96134926; PubMed=8557047;
RA Hirsch J.A., Aggarwal A.K.;
RT "Structure of the even-skipped homeodomain complexed to AT-rich DNA:
new perspectives on homeodomain specificity.";
RL EMBO J. 14:6280-6291(1995).
CC -!- FUNCTION: MAY PLAY A ROLE IN DETERMINING NEURONAL IDENTITY. MAY BE
CC DIRECTLY INVOLVED IN SPECIFYING IDENTITY OF INDIVIDUAL NEURONS.
CC PAIR-RULE PROTEIN REQUIRED FOR SEGMENTATION; INVOLVED IN
CC TRANSFORMING THE BROAD, SPATIAL, APERIODIC EXPRESSION PATTERNS OF
CC THE GAP GENES INTO A SYSTEM OF PRECISE PERIODIC EXPRESSION
CC PATTERNS OF THE PAIR-RULE AND SEGMENTARY POLARITY GENES.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE EVEN-SKIPPED FAMILY OF HOMEBOX
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14767; AAA28522.1; -
DR EMBL; X05138; CAA28784.1; -
DR EMBL; X05138; CAA28785.1; ALT_SEQ.
DR EMBL; AE003831; AAF58865.1; -
DR EMBL; U32087; AAB05358.1; -
DR EMBL; U32088; AAB05359.1; -
DR EMBL; U32089; AAB05360.1; -
DR EMBL; U32090; AAB05361.1; -
DR EMBL; U32091; AAB05362.1; -
DR PIR; A26066; A26066.
DR HSSP; P02833; 1SAN.
DR TRANSFAC; T00272; -
DR Flybase; FBgn0000606; eve.
DR InterPro; IPR001356; -
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PROSITE; PS00027; HOMEBOX_1; 1.

DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Developmental protein; Embryo; Pair-rule protein;
KW Transcription regulation; Homeobox; Nuclear protein.
FT DNA_BIND 70 129 HOMEBOX.
FT DOMAIN 165 179 ALA-RICH.
FT CONFLICT 300 300 L -> V (IN REF. 1).
SQ SEQUENCE 376 AA; 39970 MW; 59058E0634B1BED0 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
| | | | |
DB 274 GVGVP 278

RESULT 12
IMDH_BORBU
ID IMDH_BORBU STANDARD; PRT; 404 AA.
AC P49058;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE (EC 1.1.1.205) (IMP
DE DEHYDROGENASE) (IMPDH) (IMPD).
GN GUAB OR BBB17.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid cp26 (circular 26 kb).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE=95050198; PubMed=7961392;
RA Margolis N., Hogan D., Tilly K., Rosa P.;
RT "Plasmid location of Borrelia purine biosynthesis gene homologs.";
RL J. Bacteriol. 176:6427-6432(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,
RA Utterback T., Wattley L., McDonald L., Artiach P., Bowman J.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
burgdorferi.";
RL Nature 390:580-586(1997).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=20222989; PubMed=10758003;
RA McMillan F.M., Cahoon M., White A., Hedstrom L., Petsko G.A.,
RA Ringe D.;
RT "Crystal structure at 2.4-A resolution of Borrelia burgdorferi inosine
5'-monophosphate dehydrogenase: evidence of a substrate-induced
hinged-lid motion by loop 6.";
RT Biochemistry 39:4533-4542(2000).
CC -!- CATALYTIC ACTIVITY: INOSINE 5'-PHOSPHATE + NAD(+) + H(2)O =
CC XANTHOSINE 5'-PHOSPHATE + NADH.
CC -!- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.
CC -!- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO
CC GMP REDUCTASE.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC Modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: U13372; AAF81640.1; -;
 DR EMBL: AE000792; AAC66314.1; -;
 DR TIGR: BB17; -;
 DR PDB: 1EEP; 29-MAR-00.
 DR InterPro: IPR001093; -;
 DR Pfam: PF00478; IMPDH_C; 1.
 DR Pfam: PF01574; IMPDH_N; 1.
 DR PROSITE: PS00487; IMP_DH_GMP_RED; 1.
 KW Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis; Plasmid;
 3D-structure.
 FT BINDING 229 229 IMP.
 SQ SEQUENCE 404 AA; 43767 MW; A91D6D6C5CE522F1 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 404;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5

Db 236 GVGVP 240

RESULT 13
 ME12_HUMAN
 ID ME12_HUMAN STANDARD; PRT; 477 AA.
 AC O14770; Q9NRS3; Q9NRS2; Q9NRS1;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HOMEBOX PROTEIN MEIS2 (MEIS1-RELATED PROTEIN 1).
 GN MEIS2 OR MRG1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS MEIS2C; MEIS2D AND MEIS2E).
 RC TISSUE=BRAIN;
 RX Medline=10764806;
 RA Yang Y., Hwang C.K., D'Souza U.M., Lee S.-H., Junn E., Mouradian M.M.;
 RT "Three-amino acid extension loop homeodomain proteins Meis2 and Tcf1
 RT differentially regulate transcription.";
 RL J. Biol. Chem. 275:20734-20741(2000).
 RN [2]
 RP SEQUENCE OF 271-477 FROM N.A. (ISOFORM MEIS2B).
 RX MEDLINE=98051942; PubMed=9383298;
 RA Smith J.E. Jr., Afonja O., Yee H.T., Inghirami G., Takeshita K.;
 RT "Chromosome mapping to 15q14 and expression analysis of the human
 RT Meis2 homeobox gene.";
 RL Mamm. Genome 8:951-952(1997).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS: MEIS2A, MEIS2B, MEIS2C (SHOWN
 CC HERE), MEIS2D AND MEIS2E; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN VARIOUS TISSUES. IN HEMATOPOIETIC
 CC TISSUES, THE LYMPHOID ORGANS EXPRESS HIGH LEVELS OF MEIS2. ALSO
 CC EXPRESSED IN SOME REGIONS OF THE BRAIN, SUCH AS THE PUTAMEN.
 CC -1- SIMILARITY: BELONGS TO THE TALE/MEIS FAMILY OF HOMEBOX PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF179897; AAF81640.1; -;
 DR EMBL: AF179898; AAF81641.1; -;

DR EMBL: AF179899; AAF81642.1; -;
 DR EMBL: AF017418; AAB70270.1; -;
 DR MIN: 601740; -;
 DR InterPro: IPR001356; -;
 DR PROSITE: PS00027; HOMEBOX_1; FALSE_NEG.
 DR PROSITE: PS00071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Nuclear protein; Alternative splicing.
 FT DOMAIN 195 244 SER/THR-RICH (ACIDIC).
 FT DOMAIN 246 273 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 266 273 POLY-ASP.
 FT DNA_BIND 276 338 HOMEBOX (TALE-TYPE).
 FT VARSPLIC 346 352 MISSING (IN ISOFORM MEIS2B AND ISOFORM
 MEIS2D).
 FT VARSPLIC 301 302 HP -> VY (IN ISOFORM MEIS2E).
 FT VARSPLIC 303 477 MISSING (IN ISOFORM MEIS2E).
 FT VARSPLIC 384 401 LQSPGDYVSGGPMGMS -> PMSGMGMNMGDQWHYH
 (IN ISOFORM MEIS2A AND ISOFORM MEIS2B).
 FT VARSPLIC 402 477 MISSING (IN ISOFORM MEIS2A AND ISOFORM
 MEIS2B).
 SQ SEQUENCE 477 AA; 51790 MW; 94EBD0801A312B24 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 477;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5

Db 16 GVGVP 20

RESULT 14

ME12_MOUSE

ID ME12_MOUSE STANDARD; PRT; 477 AA.
 AC P97367; O35676; P97403; O35677; P97404;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HOMEBOX PROTEIN MEIS2 (MEIS1-RELATED PROTEIN 1).
 GN MEIS2 OR MRG1 OR STRA10.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM MEIS2A).
 RC STRAIN=SWISS WEBSTER;
 RX MEDLINE=97108670; PubMed=8950991;
 RA Nakamura T., Jenkins N.A., Copeland N.G.;
 RT "Identification of a new family of Pbx-related homeobox genes.";
 RL Oncogene 13:2235-2242(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS MEIS2A; MEIS2B; MEIS2C AND MEIS2D).
 RX MEDLINE=97477074; PubMed=9337137;
 RA Oulad-Abdelghani M., Chazaud C., Bouillet P., Sapin V., Chambon P.,
 RA Dolle P.;
 RT "Meis2, a novel mouse Pbx-related homeobox gene induced by retinoic
 RT acid during differentiation of P19 embryonal carcinoma cells.";
 RL Dev. Dyn. 210:173-183(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS MEIS2B/MRG1A AND MEIS2D/MRG1B).
 RX MEDLINE=97202105; PubMed=9049632;
 RA Steelman S., Moskow J.J., Muzynski K., North C., Druck T.,
 RA Montgomery J.C., Huebner K., Daar I.O., Buchberg A.M.;
 RT "Identification of a conserved family of Meis1-related homeobox
 RT genes.";
 RL Genome Res. 7:142-156(1997).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS: MEIS2A,
 CC MEIS2B/MRG1A, MEIS2C (SHOWN HERE) AND MEIS2D/MRG1B; ARE PRODUCED
 CC BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: DISPLAYS SPATIALLY RESTRICTED EXPRESSION
 CC PATTERNS IN THE DEVELOPING NERVOUS SYSTEM, LIMBS, FACE, AND IN

```
CC VARIOUS VISCERA. IN ADULT, IT IS MAINLY EXPRESSED IN THE BRAIN AND
CC FEMALE GENITAL TRACT, WITH A DIFFERENT DISTRIBUTION OF THE
CC ALTERNATIVE SPLICING FORMS IN THESE ORGANS. LOWER EXPRESSION IN LUNG
CC AND ONLY BASAL LEVEL IN HEART, LIVER, KIDNEY, SPLEEN, AND TESTIS.
CC -|- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN ALL STAGES OF
CC EMBRYONIC DEVELOPMENT ANALYSED (7 DAYS TO 17 DAYS).
CC -|- INDUCTION: BY RETINOIC ACID.
CC -|- SIMILARITY: BELONGS TO THE TALE/MEIS FAMILY OF HOMEOBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U57343; AAC52948.1; -
CC EMBL; AJ000504; CAA04138.1; -
CC EMBL; AJ000505; CAA04139.1; -
CC EMBL; AJ000506; CAA04140.1; -
CC EMBL; AJ000507; CAA04141.1; -
CC EMBL; U68383; AAB19193.1; -
CC EMBL; U68384; AAB19194.1; -
CC MGD; MGI:108564; Mrgl.
CC InterPro; IPR001356; -
CC PROSITE; PS00027; HOMEOBOX_1; FALSE_NEG.
CC PROSITE; PS00071; HOMEOBOX_2; 1.
CC DNA-binding; Nuclear protein; Homeobox; Alternative splicing.
CC KW DOMAIN 195 244 SER/THR-RICH.
CC FT DOMAIN 246 273 ASP/GLU-RICH (ACIDIC).
CC FT DOMAIN 266 273 POLY-ASP.
CC FT DNA_BIND 276 338 HOMEOBOX (TALE-TYPE).
CC FT VARSPLIC 346 352 MISSING (IN ISOFORM MEIS2B AND ISOFORM
CC MEIS2D).
CC FT VARSPLIC 384 401 LQSMPEGVYSQSGPMGMG -> PMSGMGMNMGDMGQMHYH
CC (IN ISOFORM MEIS2A AND ISOFORM MEIS2B).
CC FT VARSPLIC 402 477 MISSING (IN ISOFORM MEIS2A AND ISOFORM
CC MEIS2B).
CC FT CONFLICT 2 2 A -> E (IN REF. 1).
CC FT CONFLICT 316 316 T -> A (IN REF. 1).
CC SEQUENCE 477 AA; 51728 MW; 923610800B647A21 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 477;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
DB 16 GVGVP 20

RESULT 15
IMDH_PYREFU
ID IMDH_PYREFU STANDARD; PRT; 485 AA.
AC P42851.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE (EC 1.1.1.205) (IMP
DE DEHYDROGENASE) (IMPDH) (IMPD).
GN GUAB.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97045814; PubMed=8890736;
RA Collart F.R., Osipuk J., Trent J., Olsen G.J., Huberman E.;
RT "Cloning, characterization and sequence comparison of the gene coding
RT for IMP dehydrogenase from Pyrococcus furiosus.";
RL Gene 174:209-216(1996).
```

```
CC -|- FUNCTION: IMP IS THE RATE LIMITING ENZYME IN THE DE NOVO SYNTHESIS
CC OF GUANINE NUCLEOTIDES AND THEREFORE IS INVOLVED IN THE REGULATION
CC OF CELL GROWTH (BY SIMILARITY).
CC -|- CATALYTIC ACTIVITY: INOSINE 5'-PHOSPHATE + NAD(+) + H(2)O =
CC XANTHOSINE 5'-PHOSPHATE + NADH.
CC -|- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.
CC -|- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO
CC GMP REDUCTASE.
CC -|- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U08814; AAC44532.1; -
CC HSSP; P50097; IAK5.
CC InterPro; IPR000644; -
CC InterPro; IPR001093; -
CC Pfam; PF00571; CBS; 2.
CC Pfam; PF00478; IMPDH_C; 1.
CC Pfam; PF01574; IMPDH_N; 1.
CC DR PROSITE; PS00487; IMP_DH_GMP_RED; 1.
CC KW Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis; Repeat;
CC CBS domain. 96 148 CBS 1.
CC FT DOMAIN 154 206 CBS 2.
CC FT BINDING 301 301 IMP (POTENTIAL).
CC SEQUENCE 485 AA; 52900 MW; BDA60E0D349E5B5B CRC64;

Query Match 100.0%; Score 27; DB 1; Length 485;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
DB 308 GVGVP 312

Search completed: July 25, 2001, 16:51:26
Job time: 157 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2001, 16:51:06 ; Search time 33.32 Seconds
(without alignments)
19.854 Million cell updates/sec

Title: US-09-251-638-2

Perfect score: 27
Sequence: 1 GVGVP 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	27	100.0	38	005422	Q05422 mycobacteri
2	27	100.0	88	Q9RFY0	Q9RFY0 renibacteri
3	27	100.0	88	Q9R322	Q9R322 renibacteri
4	27	100.0	95	Q9DSW9	Q9DSW9 ascovirus d
5	27	100.0	153	Q9ZJ16	Q9ZJ16 helicobacte
6	27	100.0	161	Q64778	Q64778 avian adeno
7	27	100.0	178	Q78931	P78931 schizosacch
8	27	100.0	206	Q56970	Q56970 maize strea
9	27	100.0	206	Q73468	Q73468 maize strea
10	27	100.0	206	Q73472	Q73472 maize strea
11	27	100.0	206	Q73474	Q73474 maize strea
12	27	100.0	206	Q83476	Q83476 maize strea
13	27	100.0	214	Q45681	Q45681 caenorhabdi
14	27	100.0	223	Q9KXY0	Q9KXY0 streptomyc
15	27	100.0	240	Q9DDR8	Q9DDR8 gallus gall
16	27	100.0	246	Q9EX74	Q9EX74 rhodococcus
17	27	100.0	256	Q9FSJ4	Q9FSJ4 oryza sativ
18	27	100.0	258	Q9UMF5	Q9UMF5 homo sapien
19	27	100.0	259	Q9VJV0	Q9VJV0 drosophila

20	27	100.0	264	2	Q9X647	Q9X647 aeromonas h
21	27	100.0	281	2	Q53826	Q53826 streptomyc
22	27	100.0	282	14	Q64770	Q64770 avian adeno
23	27	100.0	284	14	Q9ICW9	Q9ICW9 kalamchoe l
24	27	100.0	290	3	Q08909	Q08909 saccharomyc
25	27	100.0	292	2	Q9EY07	Q9EY07 staphylococ
26	27	100.0	298	11	Q35726	Q35726 marmota mar
27	27	100.0	303	2	P71617	P71617 mycobacteri
28	27	100.0	308	2	Q9RS31	Q9RS31 deinococcus
29	27	100.0	309	2	Q9F3F5	Q9F3F5 streptomyc
30	27	100.0	326	10	Q9FSR1	Q9FSR1 oryza sativ
31	27	100.0	336	5	Q09546	Q09546 caenorhabdi
32	27	100.0	341	2	Q9RD71	Q9RD71 streptomyc
33	27	100.0	346	1	Q58797	Q58797 methanococc
34	27	100.0	367	1	Q9HRA3	Q9HRA3 halobacteri
35	27	100.0	368	10	Q9PPA9	Q9PPA9 oryza sativ
36	27	100.0	376	5	Q9W566	Q9W566 drosophila
37	27	100.0	378	2	Q85345	Q85345 pseudomonas
38	27	100.0	378	2	Q911S4	Q911S4 pseudomonas
39	27	100.0	387	2	P73853	P73853 synecocyst
40	27	100.0	390	13	Q9PTF0	Q9PTF0 gallus gall
41	27	100.0	393	13	Q9PTM7	Q9PTM7 brachydanio
42	27	100.0	393	13	Q9DDE0	Q9DDE0 brachydanio
43	27	100.0	398	2	P72961	P72961 synecocyst
44	27	100.0	399	5	Q23129	Q23129 caenorhabdi
45	27	100.0	401	13	Q9PTH1	Q9PTH1 gallus gall

ALIGNMENTS

RESULT 1
005422 ID 005422 PRELIMINARY; PRT; 38 AA.
AC 005422;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE PHYTOENE DEHYDROGENASE (FRAGMENT).
GN CRTI.
OS Mycobacterium marinum.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M;
RX MEDLINE=97440138; PubMed=9294446;
RA Ramakrishnan L., Tran H.T., Federspiel N.A., Falkow S.;
RT "A crtB homolog essential for photochromogenicity in Mycobacterium
RT marinum: isolation, characterization, and gene disruption via
RT homologous recombination.";
RL J. Bacteriol. 179:5862-5868(1997).
DR EMBL; U92075; AAB71427.1; -;
FT NON_TER 1
SQ SEQUENCE 38 AA; 3986 MW; 6E46332707CCDCAB CRC64;

Query Match 100.0%; Score 27; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
| | | | |
DB 3 GVGVP 7

RESULT 2
Q9RFY0 ID 09RFY0 PRELIMINARY; PRT; 88 AA.
AC Q9RFY0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE HYPOTHETICAL 10.0 KDA PROTEIN.
 OS Renibacterium salmoninarum.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Micrococciaceae; Micrococcaceae; Renibacterium.
 OX NCBI_TaxID=1646;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC33209;
 RA Rhodes L.D., Strom M.S.;
 RT "First insertion sequence identified in the salmon pathogen,
 RT Renibacterium salmoninarum."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF163691; AAF22729.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 88 AA; 9971 MW; 495F4BFEE8B11FB3 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 88;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 Db 9 GVGVP 13

RESULT 3
 ID Q9R322 PRELIMINARY; PRT; 88 AA.
 AC Q9R322;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE HYPOTHETICAL 9.9 KDA PROTEIN.
 OS Renibacterium salmoninarum.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Micrococciaceae; Micrococcaceae; Renibacterium.
 OX NCBI_TaxID=1646;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC33209;
 RA Rhodes L.D., Strom M.S.;
 RT "First insertion sequence identified in the salmon pathogen,
 RT Renibacterium salmoninarum."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF163694; AAF22735.1; -
 DR EMBL; AF163688; AAF22723.1; -
 DR EMBL; AF163689; AAF22725.1; -
 DR EMBL; AF163690; AAF22727.1; -
 DR EMBL; AF163692; AAF22731.1; -
 DR EMBL; AF163693; AAF22733.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 88 AA; 9941 MW; B95F5AFEE8B11FAE CRC64;

Query Match 100.0%; Score 27; DB 2; Length 88;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 Db 9 GVGVP 13

RESULT 4
 ID Q9DSW9 PRELIMINARY; PRT; 95 AA.
 AC Q9DSW9;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 10.4 KDA PROTEIN.

OS Ascovirus DPAV4.
 OC Viruses; dsDNA viruses, no RNA stage; Ascoviridae; Ascovirus.
 OX NCBI_TaxID=113365;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20540044; PubMed=11086137;
 RA Stasiak K., Demattei M.V., Federici B.A., Bigot Y.;
 RT "Phylogenetic position of the *Diadromus pulchellus* ascovirus DNA
 RT polymerase among viruses with large double-stranded DNA genomes."
 RL J. Gen. Virol. 81:3059-3072(2000).
 DR EMBL; AJ279812; CAC19116.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 95 AA; 10441 MW; 14D90613A58F944A CRC64;

Query Match 100.0%; Score 27; DB 14; Length 95;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 Db 74 GVGVP 78

RESULT 5
 ID Q9ZJ16 PRELIMINARY; PRT; 153 AA.
 AC Q9ZJ16;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE ICEAL.
 GN ICEAL.
 OS *Helicobacter pylori* (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;
 OC *Helicobacter*.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ALASKA STRAIN 214;
 RA Raudonikienė A., Berg D.E.;
 RT "Iceal gene from *H. pylori* (Alaska strain 214)."
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF001538; AAC64496.1; -
 DR InterPro; IPR003615; -
 DR SMART; SM00507; HHHC; 1.
 SQ SEQUENCE 153 AA; 17347 MW; 38DE206C4B20B4E3 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 153;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 Db 38 GVGVP 42

RESULT 6
 ID Q64778 PRELIMINARY; PRT; 161 AA.
 AC Q64778;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE DNA FRAGMENT.
 OS Avian adenovirus gall (strain Phelps) (Fowl adenovirus 1) (CELO).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadnavirus.
 OX NCBI_TaxID=10553;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CELO;
 RA Akopian T.A., Kaverina E.N., Kruglyak V.A., Naroditsky B.S.,

RA Tikhonenko T.T.;
 RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 222864; CAA80482.1; -
 SQ SEQUENCE 161 AA; 17871 MW; DDAB9634B8EC1BF6 CRC64;

Query Match 100.0%; Score 27; DB 14; Length 161;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 |||||
 Db 57 GVGVP 61

RESULT 7
 P78931
 ID P78931 PRELIMINARY; PRT; 178 AA.
 AC P78931;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
 DE FISSION YEAST DNA FOR ISP4, PEPTIDYL-PROLYL CIS-TRANS ISOMERASE, ATP1,
 CHROMOSOME II COSMID 1228 SEQUENCE.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972H-;
 RA Kohnosu A., Niwa O., Yano M., Saitoh S., Katayama T., Nagao K.,
 Yanagida M.;
 RT "S.pombe chromosome II cosmid 1228 sequence."
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D83992; BAA12179.1; -
 SQ SEQUENCE 178 AA; 20631 MW; 57BCE30E2B4CF656 CRC64;

Query Match 100.0%; Score 27; DB 3; Length 178;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 |||||
 Db 5 GVGVP 9

RESULT 8
 O56970
 ID O56970 PRELIMINARY; PRT; 206 AA.
 AC O56970;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE 21.7 KDA PROTEIN.
 OS Maize streak virus.
 OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
 OX NCBI_TaxID=10821;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2A FROM REUNION ISLAND;
 RA Isnard M., Granier M., Frutos R., Reynaud B., Peterschmitt M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2A FROM REUNION ISLAND;
 RX MEDLINE=97049057; PubMed=8893787;
 RA Peterschmitt M., Granier M., Frutos R., Reynaud B.;
 RT "Infectivity and complete nucleotide sequence of the genome of a
 genetically distinct strain of maize streak virus from Reunion
 Island.";

RL Arch. Virol. 141:1637-1650(1996).
 DR EMBL; AJ224508; CAA12004.1; -
 SQ SEQUENCE 206 AA; 21743 MW; 553E4A41A9EF3AF4 CRC64;

Query Match 100.0%; Score 27; DB 14; Length 206;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 |||||
 Db 128 GVGVP 132

RESULT 9
 O73468
 ID O73468 PRELIMINARY; PRT; 206 AA.
 AC O73468;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE 21.7 KDA PROTEIN.
 OS Maize streak virus.
 OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
 OX NCBI_TaxID=10821;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2A;
 RX MEDLINE=97049057; PubMed=8893787;
 RA Peterschmitt M., Granier M., Frutos R., Reynaud B.;
 RT "Infectivity and complete nucleotide sequence of the genome of a
 genetically distinct strain of maize streak virus from Reunion
 Island.";
 RL Arch. Virol. 141:1637-1650(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2A;
 RA Isnard M., Granier M., Frutos R., Reynaud B., Peterschmitt M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ225006; CAA12303.1; -
 SQ SEQUENCE 206 AA; 21757 MW; 4F6F098FA9EF3AF0 CRC64;

Query Match 100.0%; Score 27; DB 14; Length 206;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 |||||
 Db 128 GVGVP 132

RESULT 10
 O73472
 ID O73472 PRELIMINARY; PRT; 206 AA.
 AC O73472;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE 21.7 KDA PROTEIN.
 OS Maize streak virus.
 OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
 OX NCBI_TaxID=10821;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP1;
 RX MEDLINE=97049057; PubMed=8893787;
 RA Peterschmitt M., Granier M., Frutos R., Reynaud B.;
 RT "Infectivity and complete nucleotide sequence of the genome of a
 genetically distinct strain of maize streak virus from Reunion
 Island.";
 RL Arch. Virol. 141:1637-1650(1996).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN-SP1;
 RA Isnard M., Granier M., Frutos R., Reynaud B., Peterschmitt M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ225007; CAAL12310.1; -
 SQ SEQUENCE 206 AA; 21842 MW; 488897F1B3459BE1 CRC64;

Query Match 100.0%; Score 27; DB 14; Length 206;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 Db 128 GVGVP 132

RESULT 11
 ID 073474 PRELIMINARY; PRT; 206 AA.
 AC 073474;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE 21.7 KDA PROTEIN.
 OS Maize streak virus.
 OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
 OX NCBI_TaxID=10821;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SP2;
 RX MEDLINE=97049057; PubMed=8893787;
 RA Peterschmitt M., Granier M., Frutos R., Reynaud B.;
 RL "Infectivity and complete nucleotide sequence of the genome of a
 RT genetically distinct strain of maize streak virus from Reunion
 RI Island.";
 RL Arch. Virol. 141:1637-1650(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SP2;
 RA Isnard M., Granier M., Frutos R., Reynaud B., Peterschmitt M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ225008; CAAL12317.1; -
 SQ SEQUENCE 206 AA; 21726 MW; 428987F1B8EE2646 CRC64;

Query Match 100.0%; Score 27; DB 14; Length 206;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 Db 128 GVGVP 132

RESULT 12
 ID 083476 PRELIMINARY; PRT; 206 AA.
 AC 083476;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE 21.7 KDA PROTEIN.
 OS Maize streak virus.
 OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
 OX NCBI_TaxID=10821;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isnard M., Granier M., Frutos R., Reynaud B., Peterschmitt M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RX MEDLINE=97049057; PubMed=8893787;

RA Peterschmitt M., Granier M., Frutos R., Reynaud B.;
 RT "Infectivity and complete nucleotide sequence of the genome of a
 RI genetically distinct strain of maize streak virus from Reunion
 RI Island.";
 RL Arch. Virol. 141:1637-1650(1996).
 DR EMBL: X94330; CAA63994.1; -
 DR EMBL: AJ224507; CAAL1997.1; -
 DR EMBL: AJ224504; CAAL1976.1; -
 DR EMBL: AJ224505; CAAL1983.1; -
 DR EMBL: AJ224506; CAAL1990.1; -
 DR EMBL: AJ225011; CAAL12338.1; -
 DR EMBL: AJ225009; CAAL12324.1; -
 DR EMBL: AJ225010; CAAL12331.1; -
 SQ SEQUENCE 206 AA; 21743 MW; 488897F1A9EF3AF0 CRC64;

Query Match 100.0%; Score 27; DB 14; Length 206;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 Db 128 GVGVP 132

RESULT 13
 ID 045681 PRELIMINARY; PRT; 214 AA.
 AC 045681;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
 DE K10H10.4 PROTEIN.
 GN K10H10.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Percy C.;
 RA Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thiermy-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Reinstock L., Wilkinson-Sproat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RI elegans.";

Query Match 100.0%; Score 27; DB 5; Length 214;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 Db 195 GVGVP 199

RESULT 14
 Q9KYY0

ID Q9KYY0 PRELIMINARY; PRT; 223 AA.
 AC Q9KYY0:
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE PUTATIVE PHOSPHATASE.
 GN SCE33.02C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Saunders D.C., Harris D.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapaitte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL355774; CAB90911.1; -;
 DR InterPro; IPR001454; -;
 DR Pfam; PF00702; Hydrolase; 1;
 DR SEQUENCE 223 AA; 23869 MW; 1D14700B7DA9B241 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 223;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 DB 202 GVGVP 206

RESULT 15
 Q9DDR8
 ID Q9DDR8 PRELIMINARY; PRT; 240 AA.
 AC Q9DDR8:
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE NA/PI COTRANSPORTER NAPI-IIA (FRAGMENT).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Werner A., Kinne R.K.H. Jr.;
 RT "The evolution of the Na/Pi cotransport systems.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF297188; AAG35802.1; -;
 FT NON_TER 1
 FT NON_TER 240 240
 SEQUENCE 240 AA; 25214 MW; F02283NAC242496C CRC64;

Query Match 100.0%; Score 27; DB 13; Length 240;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
 DB 118 GVGVP 122

Search completed: July 25, 2001, 16:51:08
 Job time: 139 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2001, 16:49:30 ; Search time 32.96 Seconds
(without alignments)
9.197 Million cell updates/sec

Title: US-09-251-638-2

Perfect score: 27

Sequence: 1 GVGVP 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	27	100.0	5 11 AAR04173	Monomer of repetat
2	27	100.0	5 16 AAR65230	Fibrinogen/beta-ca
3	27	100.0	5 17 AAW01440	Beta turn peptide
4	27	100.0	5 17 AAW09207	Elastin repeat pep
5	27	100.0	5 18 AAW22711	Bioelastomeric rep
6	27	100.0	5 18 AAW26331	Elastin repeat uni
7	27	100.0	5 18 AAW18263	Transglutaminase c
8	27	100.0	5 19 AAW53505	Peptide reapeat un
9	27	100.0	5 20 AAY31689	Pentapeptide used
10	27	100.0	5 20 AAY04125	Protein based poly
11	27	100.0	5 21 AAY51876	Elastin peptide re

12	27	100.0	5 21 AAY78245	Repeating unit SEQ
13	27	100.0	7 20 AAY39865	Elastin peptide fr
14	27	100.0	7 21 AAY83446	Elastin derived pe
15	27	100.0	8 20 AAY06866	Short elastin pept
16	27	100.0	9 20 AAY06867	Short elastin pept
17	27	100.0	13 10 AAY91307	Sequence of beta-t
18	27	100.0	15 20 AAY55862	Apolipoprotein fra
19	27	100.0	16 19 AAW47350	Inhibitor of LDL b
20	27	100.0	16 20 AAY55879	Apolipoprotein fra
21	27	100.0	16 20 AAY33104	Elastin derived re
22	27	100.0	18 19 AAW44902	Bovine "polyprollin
23	27	100.0	18 20 AAY06871	Short elastin pept
24	27	100.0	19 20 AAY06872	Short elastin pept
25	27	100.0	20 13 AAR26383	Sequence of peptid
26	27	100.0	20 14 AAR40000	Scintigraph imagin
27	27	100.0	20 17 AAW01445	Bioelastic polymer
28	27	100.0	20 17 AAW11073	Atherosclerotic pl
29	27	100.0	20 18 AAW31123	Atherosclerotic pl
30	27	100.0	20 20 AAY55861	Apolipoprotein fra
31	27	100.0	20 20 AAY06873	Short elastin pept
32	27	100.0	20 21 AAY80347	N-terminal fusion
33	27	100.0	21 12 AAR15136	Vascular injury af
34	27	100.0	21 14 AAR42543	Leukocyte-binding
35	27	100.0	21 19 AAW47336	Inhibitor of LDL b
36	27	100.0	21 20 AAY55875	Apolipoprotein fra
37	27	100.0	21 20 AAY33100	Elastin derived pe
38	27	100.0	21 20 AAY06869	Short elastin pept
39	27	100.0	21 20 AAY06870	Short elastin pept
40	27	100.0	21 20 AAY06868	Short elastin pept
41	27	100.0	22 12 AAR14949	Medial middle laye
42	27	100.0	22 14 AAR42536	Leukocyte-binding
43	27	100.0	22 17 AAW09215	Cleavage site sequ
44	27	100.0	22 19 AAW53542	Amino acid sequenc
45	27	100.0	22 21 AAY51885	Crosslinked protei

ALIGNMENTS

RESULT 1	
AAAR04173	
ID AAR04173 standard; protein; 5 AA.	
XX	
AC AAR04173;	
XX	
DT 10-SEP-1990 (first entry)	
XX	
DE Monomer of repetitive primary structure of a synthetic fibrous	
DE protein.	
XX	
KW Collagen; elastin; keratin; troponin C; silk; dopa protein;	
KW synthetic skin; cosmetics; bioadhesive.	
XX	
OS Synthetic.	
XX	
PN WO9003438-A.	
XX	
PD 05-APR-1990.	
XX	
PF 06-SEP-1989; 89WO-0003839.	
XX	
PR 30-SEP-1988; 88US-0251714.	
XX	
PA (ALLC) ALLIED SIGNAL INC.	
XX	
PI Goldberg I, Salerno AJ;	
XX	
DR WPI; 1990-132274/17.	
XX	
XX New bacterial strains for heterologous gene expression -	
PT contg. elements for initiating activity, retarding proteolysis	
PT and stabilising heterologous genes.	
XX	

PS Disclosure; ; lpp; English.

CC The polymer of the peptide may act as an analogue to peptides
 CC contg. a repetitive or quasi-repetitive structure eg. collagen,
 CC elastin, keratin, tropoin C, dopa proteins, silk proteins,
 CC bioadhesive proteins and insect cuticle proteins. These products
 CC may be useful in the manufacture of fibrous products, synthetic
 CC skin and cosmetic additives.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 11; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5 0; Indels 0; Gaps 0;

Db 1 gvgvp 5

RESULT 2

AAR65230
 ID AAR65230 standard; peptide; 5 AA.

AC AAR65230;

XX 12-OCT-1995 (first entry)

XX Fibrinogen/beta-casein transglutaminase cross-linkable copolymer.

XX Transglutaminase cross-linkable copolymer; fibrinogen; beta-casein;
 KW biocompatible; bioadhesive polymers; wound repair; prosthetics;
 KW bone and soft tissue matrices; controlled drug release carriers.

XX Synthetic.

XX WO9505396-A.

XX 23-FEB-1995.

XX 05-AUG-1994; 94WO-US08754.

XX 13-AUG-1993; 93US-0106509.

XX (ZYMO) ZYMOGENETICS INC.

XX Busby SJ, Labroo VM;

XX WPI; 1995-098722/13.

XX New peptide(s) based on fibrinogen and beta-casein -
 PT cross-linkable by transglutaminase, used for preparing
 PT biocompatible, bioadhesive polymers

XX Claim 16; Page 48; 58pp; English.

XX A biocompatible, bioadhesive, transglutaminase cross-linkable
 CC copolymer comprises a first polypeptide monomer from 13-120 amino
 CC acids containing a segment of the formula S1-Y-S2 cross-linkable
 CC by a transglutaminase, where S1 = AAR65218, S2 = AAR65219 and Y = a
 CC spacer of 0-7 amino acids pref. AAR65220/21/22 or AAR65223, and
 CC a second polypeptide monomer selected from AAR65228-R65231 (the
 CC first monomer can also be AAR65227). The copolymer can be used to
 CC produce tissue adhesives, wound repair formulations, rigid
 CC prosthetics, matrices for the replacement of bone and soft tissue
 CC structures and carriers for controlled drug release compns.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 16; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5

Db 1 gvgvp 5

RESULT 3

AAW01440
 ID AAW01440 standard; peptide; 5 AA.

XX AAW01440;

XX 18-JUN-1997 (first entry)

XX Beta turn peptide GVGVP.

XX Beta turn; artificial polymer; inverse temperature transition;
 KW bioelastomer; repeat unit.

XX Synthetic.

XX WO9632406-A1.

XX 17-OCT-1996.

XX 15-APR-1996; 96WO-US05186.

XX 13-OCT-1995; 95US-0543020.

XX 14-APR-1995; 95US-0423516.

XX (BIOE-) BIOELASTICS RES LTD.

XX (UABR-) UAB RES FOUND.

XX McPherson DT, Urry DW, Xu J;

XX WPI; 1996-477065/47.

XX Purifying artificial polymer exhibiting reversible inverse temp.
 PT transition - specifically bioelastomeric polymers from complex mixtures
 PT of proteins, e.g., those found after expression in host.

XX Example 1; Page 27; 46pp; English.

XX This sequence represents a beta turn peptide. This sequence was used as
 CC a repeating unit in an artificial polymer, which exhibits reversible
 CC inverse temperature transition, purified by the method of the invention.
 CC The method of the invention comprises dissolving the polymer in an
 CC aqueous medium, and adjusting the relative temperature of the aqueous
 CC medium to the effective transition temperature of the polymer, so that
 CC the temperature of the aqueous medium is below the effective transition
 CC temperature. Any particulate material present in the medium is then
 CC removed, and the relative temperature of the aqueous medium is adjusted
 CC to the effective transition temperature of the polymer, so that the
 CC temperature of the aqueous medium is above the effective transition
 CC temperature of the polymer. Dense phase material comprising the polymer
 CC is then collected from the medium. Any of these steps can be repeated
 CC until a desired level of purity is reached. The method is used for the
 CC purification of polymers that exhibit inverse temperature transition to a
 CC high degree of purity for their commercial application, specifically
 CC bioelastomers irrespective of their source. The method is particularly
 CC advantageous for purifying bioelastomeric polypeptides from complex
 CC mixtures of proteins, e.g. those found after expression of bioelastomeric
 CC polypeptides in a host, and can also be used to purify chemically
 CC synthesised peptides.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 17; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYGVP 5
|||||
Db 1 gvgvp 5

RESULT 4

AAW09207
ID AAW09207 standard; peptide; 5 AA.
AC AAW09207;
XX
XX 29-MAR-1997 (first entry)
XX Elastin repeat peptide.
DE
XX
XX Elastin; repeat unit; protein polymer; crosslinking;
KW adhesive; tissue sealant; biocompatible film.
XX
XX Synthetic.
XX W09634618-A1.
XX
XX 07-NOV-1996.
XX
XX 02-MAY-1996; 96WO-US06229.
XX
XX 05-MAY-1995; 95US-0435641.
XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
XX
XX Cappelletto J, Stedronsky ER;
PI
XX
XX WPI; 1996-505895/50.

XX Tissue repair adhesive comprising polymer of structural protein
XX repeat units - contg. hetero-atom functional gps. reactive with
XX crosslinking agent, combines biocompatibility and high bonding
XX strength.
XX
XX Claim 4; Page 96; 103pp; English.

XX This peptide is a repeat unit derived from elastin, and may be
XX used in construction of a protein polymer backbone, which may be
XX crosslinked chemically to form a strongly adherent tissue adhesive
XX or sealant. The polymer may be incubated with a polyfunctional
XX crosslinking agent to form the adhesive. The repeat unit may be
XX used with silk fibroin peptide AAW09206 at a ratio of 1:16:1 to form a
XX block copolymer protein, e.g. AAW09213, AAW09218, AAW09219, AAW09221 or
XX AAW09223. At least one of the repeats may be modified to introduce a
XX crosslinking functionality. The polymer may be used to seal
XX defects in vessel walls, e.g. artery, vein, capillary, lung, dura or
XX colon, to increase tissue mass, or to produce biocompatible films
XX for in vivo use. The sealants have the biocompatibility of fibrin
XX glues, but set more quickly, and have greater strength. They are
XX readily prepared by recombinant methods, are easy to administer,
XX and are gradually resorbed.

Sequence 5 AA;

Query Match 100.0%; Score 27; DB 17; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYGVP 5
|||||
Db 1 gvgvp 5

RESULT 5

AAW22711
ID AAW22711 standard; peptide; 5 AA.
XX

AC AAW22711;

XX
XX 26-FEB-1998 (first entry)

XX Bioelastomeric repeating unit 4 responsive to electrical energy.
DE
XX
XX Bioelastomeric repeating unit; polymer; inverse temperature transition;
KW electrical energy change; polarity; hydrophobicity; mechanical work;
KW surgical suture; heat shrinkable membrane; desalination.
XX
XX Synthetic.

XX W09723729-A1.

XX 03-JUL-1997.

XX 07-JUN-1996; 96WO-US09776.

XX 07-JUN-1995; 95US-0487594.

XX (URRY/) URRY D W.

XX Urry DW;

XX WPI; 1997-363360/33.

XX Bioelastic polymer responsive to electrical energy - comprising
XX beta turn and residue(s) with side chain that changes polarity or
XX hydrophobicity in response to electrical energy change, useful for
XX mechanical work or light stimulated contraction
XX
XX Example 1; Page 47; 60pp; English.

XX AAW22711-16 represent bioelastomeric repeating units that are used to
XX create a polymer responsive to electrical energy. This bioelastic
XX polymer comprises a bioelastomeric polypeptide having an inverse
XX temperature transition, and a repeating bioelastomeric unit containing
XX at least 1 beta-turn and residues with a side chain. The side chain of
XX these residues responds to an electrical energy change by altering its
XX polarity or hydrophobicity. The side chain is present in sufficient
XX amounts to provide a shift in the inverse temperature transition of the
XX polymer upon the change in exposure to electrical energy. The polymer may
XX also contain a second amino acid with a side chain capable of undergoing
XX a change in an aqueous environment. The polymers (AAW34881-83) consist of
XX the formulas described in AAW22716 and AAW22714 and below:

XX poly[0.57(GVGVP), 0.33(GRGVP), 0.1(GKGV)]
XX The expansion or contraction of the polymer upon exposure to a change
XX in electrical energy can be used to produce mechanical work when the
XX polymer is constrained. It can also cause turbidity and chemical changes
XX in a closed environment. The polymer can be used in surgical sutures,
XX especially for microsurgery, heat shrinkable membranes, controlled
XX herbicide or pesticide release or for desalination.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 18; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
|||||
Db 1 gvgvp 5

RESULT 6

AAW26331
ID AAW26331 standard; Peptide; 5 AA.

XX AAW26331;

XX 19-NOV-1997 (first entry)
XX

DE Elastin repeat unit.

XX Elastin; protein polymer.
 XX Synthetic.
 XX US5641648-A.
 XX 24-JUN-1997.
 XX 04-NOV-1986; 86US-0927258.
 XX 29-DEC-1993; 93US-0175155.
 XX 04-NOV-1986; 86US-0927258.
 XX 29-OCT-1987; 87US-0114618.
 XX 09-NOV-1988; 88US-0269429.
 XX 08-NOV-1990; 90US-0609716.
 XX 22-APR-1993; 93US-0053049.

XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
 XX Cappello J, Ferrari FA, Richardson C;
 XX WPI; 1997-340943/31.
 XX Preparation of synthetic DNA encoding a protein comprising short
 PT repeats - by synthesising oligomers, annealing and oligomerising
 PT these, particularly to produce proteins that mimic silk, collagen
 PT etc
 XX Disclosure; Column 12; 90pp; English.
 XX This peptide represents a repeat unit found in elastin. Methods
 CC are claimed for preparing protein polymers that contain repeating
 CC units based on those found in elastin, fibroin, keratin, etc. (see
 CC AAW26329, AAW26332, AAW26334 and AAW26336). These involve: synthesising
 CC pairs of single-stranded oligomers, each member of the pair
 CC overlapping except at the protruding ends; hybridising each pair to
 CC double-stranded segments; combining the double-stranded segments,
 CC or their cloned copies, in a cloning vector to form a monomer;
 CC excising the monomer from the vector; and oligomerising the monomer
 CC to produce a multimer. Such polymers comprise fibrous or
 CC structural proteins, including crystalline, elastomeric, tough and
 CC bony materials such as those that mimic (and can substitute for)
 CC silk, elastin, collagen, keratin etc. The properties of the protein
 CC can be controlled by varying the type of units in the monomer, the
 CC number of units per multimer, the spacing between them and the
 CC number of multimer repeats.

PS Sequence 5 AA;

XX Query Match 100.0%; Score 27; DB 18; Length 5;
 XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 DB 1 gvgvp 5
 RESULT 7
 AAW18263
 ID AAW18263 standard; peptide; 5 AA.
 XX AC AAW18263;
 XX 29-AUG-1997 (first entry)
 XX Transglutaminase cross-linkable polypeptide elastomeric peptide.
 DE Elastomeric; homopolymer; copolymer; tissue sealant; skin graft;
 KW burn.

Query Match 100.0%; Score 27; DB 18; Length 5;

XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 DB 1 gvgvp 5

RESULT 7

AAW18263
 ID AAW18263 standard; peptide; 5 AA.
 XX AC AAW18263;
 XX 29-AUG-1997 (first entry)
 XX Transglutaminase cross-linkable polypeptide elastomeric peptide.
 DE Elastomeric; homopolymer; copolymer; tissue sealant; skin graft;
 KW burn.

XX Synthetic.
 XX WO9640780-A1.
 XX 19-DEC-1996.
 XX 31-MAY-1996; 96WO-US08269.
 XX 07-JUN-1995; 95US-0483236.
 XX (ZYMO) ZYMOGENETICS INC.
 XX Busby SJ, Labroo VM;
 XX WPI; 1997-052237/05.

XX Trans:glutaminase cross-linkable peptide(s) - used in the mfr. of
 PT biocompatible, bio:adhesive tissue sealant and wound healing
 PT preparations.
 XX Claim 6; Page 48; 55pp; English.
 XX Novel polypeptides optionally having one or both of the amino-terminus
 CC and carboxy-terminus flanked by an elastomeric peptide, which is cross-
 CC linkable by a transglutaminase comprises a segment of formula:
 CC SI-Y-S2. The present sequence represents a specifically claimed
 CC example of an elastomeric peptide. The homo- and copolymers produced
 CC are useful in tissue sealant and wound healing formulations. Tissue
 CC sealants are useful in skin grafting for burn victims and for sealing
 CC surgical and other wounds.

XX Sequence 5 AA;
 XX Query Match 100.0%; Score 27; DB 18; Length 5;
 XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 DB 1 gvgvp 5

RESULT 8

AAW53505
 ID AAW53505 standard; Protein; 5 AA.
 XX AC AAW53505;
 XX 10-AUG-1998 (first entry)
 XX Peptide repeat unit 3.
 XX Peptide repeat unit; DNA repeat unit; high molecular weight polymer;
 KW synthetic silk; silk worm.

Query Match 100.0%; Score 27; DB 18; Length 5;

XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 DB 1 gvgvp 5
 RESULT 8
 AAW53505
 ID AAW53505 standard; Protein; 5 AA.
 XX AC AAW53505;
 XX 10-AUG-1998 (first entry)
 XX Peptide repeat unit 3.
 XX Peptide repeat unit; DNA repeat unit; high molecular weight polymer;
 KW synthetic silk; silk worm.

Query Match 100.0%; Score 27; DB 18; Length 5;

XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 DB 1 gvgvp 5

RESULT 8

AAW53505
 ID AAW53505 standard; Protein; 5 AA.
 XX AC AAW53505;
 XX 10-AUG-1998 (first entry)
 XX Peptide repeat unit 3.
 XX Peptide repeat unit; DNA repeat unit; high molecular weight polymer;
 KW synthetic silk; silk worm.

Query Match 100.0%; Score 27; DB 18; Length 5;

XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 DB 1 gvgvp 5
 RESULT 8
 AAW53505
 ID AAW53505 standard; Protein; 5 AA.
 XX AC AAW53505;
 XX 10-AUG-1998 (first entry)
 XX Peptide repeat unit 3.
 XX Peptide repeat unit; DNA repeat unit; high molecular weight polymer;
 KW synthetic silk; silk worm.

Query Match 100.0%; Score 27; DB 18; Length 5;

XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 DB 1 gvgvp 5
 RESULT 8
 AAW53505
 ID AAW53505 standard; Protein; 5 AA.
 XX AC AAW53505;
 XX 10-AUG-1998 (first entry)
 XX Peptide repeat unit 3.
 XX Peptide repeat unit; DNA repeat unit; high molecular weight polymer;
 KW synthetic silk; silk worm.

Query Match 100.0%; Score 27; DB 18; Length 5;

XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 DB 1 gvgvp 5
 RESULT 8
 AAW53505
 ID AAW53505 standard; Protein; 5 AA.
 XX AC AAW53505;
 XX 10-AUG-1998 (first entry)
 XX Peptide repeat unit 3.
 XX Peptide repeat unit; DNA repeat unit; high molecular weight polymer;
 KW synthetic silk; silk worm.

Query Match 100.0%; Score 27; DB 18; Length 5;

XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVGVP 5
 DB 1 gvgvp 5
 RESULT 8
 AAW53505
 ID AAW53505 standard; Protein; 5 AA.
 XX AC AAW53505;
 XX 10-AUG-1998 (first entry)
 XX Peptide repeat unit 3.
 XX Peptide repeat unit; DNA repeat unit; high molecular weight polymer;
 KW synthetic silk; silk worm.

XX Preparation of synthetic repetitive DNA - useful for construction of
PT large protein polymers having repeating units, used in structural
PT material, e.g. synthetic silk
XX
PS Claim 11; Page 22; 127pp; English.
XX
CC This is the amino acid sequence of the peptide repeat unit, used in the
CC method of the invention, which involves the preparation of synthetic
CC DNA sequence having repeating units from about 3-15 codons and encoding
CC a protein of at least about 30 kDa. The method is useful for the
CC production of high molecular weight polymers (e.g. synthetic silk),
CC either nucleic acids or peptides that are the expression products of
CC the nucleic acids and particularly high molecular weight peptides
CC containing repeating units which are useful as structural materials.
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 27; DB 19; Length 5;
Best Local Similarity 100.0%; Pred. NO. 3.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
Db 1 gvgvp 5
|||||

RESULT 9
AAV31689
ID AAY31689 standard; Peptide; 5 AA.
XX
AC AAY31689;
XX
DT 22-NOV-1999 (first entry)
XX
DE Pentapeptide used in novel elastomer polymers.
XX
KW Elastomer; bioelastomer; polymer; tissue augmentation; elastin;
KW tissue restoration; tissue reconstruction; tissue repair; implant.
XX
OS Synthetic.
XX
PN WO9943271-AL.
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04440.
XX
PR 29-MAY-1998; 98US-0087155.
PR 27-FEB-1998; 98US-0076297.
XX
PA (BIOE-) BIOELASTICS RES LTD.
XX
PI Glazer PA, Parker TM, Urry DW;
XX
DR WPI; 1999-540487/45.
XX
PT Augmentation or restoration of mammalian tissue by injecting
PT solution of peptide polymer, used for soft or hard tissue
PT reconstruction, especially of intervertebral disks
XX
PS Claim 9; Page 81; 133pp; English.
XX
CC The present sequence represents an example of a pentapeptide that
CC is used in novel bioelastic polymers. The invention provides a
CC method of tissue augmentation by injecting a polymer comprising
CC repeating peptide monomeric units selected from nonapeptide,
CC pentapeptide and tetrapeptide monomeric units, where the monomeric
CC units form a series of beta-turns separated by dynamic bridging
CC segments. The polymer has an inverse temperature transition value
CC that is less than the tissue temp. and is injected in water
CC solution at coacervate concentration. The polymer can be injected

CC at periurethral or subdermal sites (for treatment of urinary
CC incontinence or for cosmetic purposes), or into hard or soft
CC tissue, e.g. for repair of traumatic injury. A specific
CC application is restoration of intervertebral discs. The elastin
CC type polymer based on the present pentapeptide, when prepared with
CC repeat units in the order of 200 and when cross-linked by
CC gamma-irradiation, forms an elastic matrix with an elastic modulus
CC approximating that of the vascular wall, providing the capacity
CC to match compliance over a wide range of biological tissues.
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 27; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. NO. 3.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
Db 1 gvgvp 5
|||||

RESULT 10
AAV04125
ID AAY04125 standard; peptide; 5 AA.
XX
AC AAY04125;
XX
DT 11-JUN-1999 (first entry)
XX
DE Protein based polymer pentamer peptide sequence.
XX
KW Chloroplast integration; expression vector; transformation; PBP;
KW protein based polymer; herbicide resistance.
XX
OS Synthetic.
XX
PN WO9910513-AL.
XX
PD 04-MAR-1999.
XX
PF 05-AUG-1998; 98WO-IB01199.
XX
PR 15-MAY-1998; 98US-0079640.
PR 07-AUG-1997; 97US-0055314.
PR 23-MAR-1998; 98US-0079042.
XX
PA (UYAU-) UNIV AUBURN.
XX
PI Daniell H;
XX
DR WPI; 1999-190626/16.
XX
PT New universal chloroplast integration and expression vectors -
PT containing an expression cassette for a molecule of interest and
PT flanking sequences for insertion into the chloroplast genome by
PT homologous recombination
XX
PS Claim 109; Page 81; 131pp; English.
XX
CC The present invention describes a novel universal integration and
CC expression vector (A) competent for stably transforming the chloroplast
CC genome of different plant species. Products from the present invention
CC can be used for producing transformed plants such as maize, rice, grass,
CC rye, barley, oat, wheat, soybean, peanut, grape, sweet potato, pea,
CC canola, tobacco, tomato or cotton. The transformed plants can be used
CC for producing polypeptides of interest, e.g. insulin or human serum
CC albumin. The plants can also be made resistant to herbicides e.g.
CC glyphosate, paraquat, diquat, atrazine, dinoseb, lenacil, metribuzine,
CC dinitrohexanediol, haloxyfop, clethodim and phenoxaprop, MCPA, 2,4-D,
CC diclofop, sulfonyleureas, imidazolinones, bromoxynil, methyl sulfuron,
CC chlorosulfuron, phosphinothricin or imazapyr. They can also be made
CC resistant to insects e.g. tobacco budworm, cottonboll worm or beet

CC armyworm. They can also be used for the production of protein-based
 CC polymers (PBPs) which can be used as hydrogels, elastomers and plastics
 CC and in medical applications including the prevention of post-surgical
 CC adhesions, tissue reconstruction and programmed drug delivery. The
 CC present sequence represents a specifically claimed PBP pentamer peptide
 CC sequence.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVGVP 5

Db 1 GVGVP 5

RESULT 11

AAV51876
 ID AAV51876 standard; peptide; 5 AA.

XX AAV51876;

AC 22-JUN-2000 (first entry)

XX Elastin peptide repeat unit.

DE Crosslinked protein; fibrin glue; elastin; tissue adhesive; sealant.

XX Unidentified.

XX US6033654-A.

XX 07-MAR-2000.

XX 02-MAY-1996; 96US-0642246.

XX 05-MAY-1995; 95US-0435641.

XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

XX Cappelletto J, Stedronsky ER;

XX WPI; 2000-255682/22.

XX Crosslinked protein composition, useful as tissue adhesive or sealant,
 PT comprises peptide repeating units that contain functional groups
 PT reactive with crosslinker.

PS Claim 1; Column 45-46; 45pp; English.

XX This invention describes a novel crosslinked protein composition (A) in
 CC which, before crosslinking, the protein (I) is new and comprises at
 CC least 70 wt. % of repeating units GAGAGS (1) and GVGVP (2), and in at
 CC least two repeating units an amino acid (aa) is substituted by Lys or
 CC Arg to provide a Lys/Arg equivalent weight of 1-20 Kd. (1) contains at
 CC least two aa having a functional group reactive with at least one of
 CC aldehyde, iso(thio)cyanate and activated carboxy. (1) have similar
 CC biocompatibility to fibrin glues, but set more quickly and give a bond
 CC with greater shear strength. They are made from readily available natural
 CC sources, are easy to administer and are gradually resorbed. This sequence
 CC represents a repeating peptide unit derived from elastin.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVGVP 5

Db 1 GVGVP 5

RESULT 12

AAV78245
 ID AAV78245 standard; Peptide; 5 AA.

XX AAV78245;

AC 03-MAY-2000 (first entry)

XX Repeating unit SEQ ID NO:23.

XX Repetitive unit; repeat; prosthetic device; fibre; silk protein;
 KW fibrous; structure; film; gel; membrane; amorphous; adhesive;
 KW coating; viscous fluid; emulsion; crystalline; elastomeric;
 KW tough; bony material; elastin; collagen; keratin.

XX Bombyx mori.

XX US6018030-A.

XX 25-JAN-2000.

XX 07-JUN-1995; 95US-0482085.

XX 29-OCT-1987; 87US-0114618.

XX 04-NOV-1986; 86US-0927258.

XX 22-APR-1993; 93US-0053049.

XX 29-DEC-1993; 93US-0175155.

XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

XX Ferrari FA, Richardson C, Causey S, Pollock TJ, Cappelletto J;
 XX Crissman JW, Chambers J;

XX WPI; 2000-146766/13.

XX Proteins containing repeat units of naturally occurring proteins,
 PT useful as fibrous or structural proteins -

PS Claim 10; Column 182; 102pp; English.

XX The present invention describes a recombinantly produced protein (I) of
 CC 30 to 250 kilodaltons (kDa), where 50% of the protein is comprised of
 CC at least 1 repeating unit of a naturally occurring protein. The proteins
 CC comprising repeating units may be used as fibrous or structural
 CC proteins, including crystalline, elastomeric, tough and bony materials
 CC e.g. proteins similar to but different from, silk, elastin, collagen,
 CC keratin or other naturally occurring structural polymers having
 CC repetitive amino acid sequence motifs. The proteins may have the
 CC capability of forming or participating in the formation of objects such as
 CC films, fibres, gels, membranes or may be amorphous, such as in adhesives,
 CC coatings, viscous fluids and emulsions. AAV78243 to AAV78295 and AAV88279
 CC to AAV88316 represent amino acid and nucleotide sequences used in the
 CC exemplification of the present invention.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVGVP 5

Db 1 GVGVP 5

RESULT 13

AAV39865
 ID AAV39865 standard; peptide; 7 AA.

```

XX AC AAY39865;
XX DT 03-DEC-1999 (first entry)
XX DE Elastin peptide fragment #39.
XX KW Elastin peptide; mammalian tissue; hypertension; coronary heart disease;
XX KW arteriosclerosis; angina; coronary thrombosis; neovascularisation;
XX KW chronic obstructive pulmonary disease; restenosis post angioplasty;
XX KW tissue turgor; tissue elasticity; angiogenesis; therapy.
XX OS Mammalia.
XX PN WO9945941-A1.
XX PD 16-SEP-1999.
XX PF 12-MAR-1999; 99WO-US05496.
XX PR 13-MAR-1998; 98US-0039308.
XX PA (MRSN-) MRS LLC.
XX PI Sandberg LB, Roos PJ, Mitts TF;
XX XX WPI; 1999-561616/47.
XX PT Elastin peptide fragment compositions useful for treating hypertension,
XX PT coronary heart disease and arteriosclerosis.
XX PS Claim 4; Page 22; 55pp; English.
XX CC This sequence represents an elastin peptide fragment that is used in the
XX CC composition of the invention, which is useful in treating a condition of
XX CC mammalian tissue. The elastin peptide fragment composition is useful for
XX CC treating hypertension, coronary heart disease, arteriosclerosis, angina,
XX CC coronary thrombosis, chronic obstructive pulmonary disease and restenosis
XX CC post angioplasty. The mammalian tissue being treated is a mammalian blood
XX CC vessel. The composition is useful in improving tissue turgor. The
XX CC composition can be used in a method for enhancing tissue elasticity. The
XX CC method can be used to enhance tissue, especially skin, elasticity.
XX CC Increased elasticity and/or physical appearance of the skin gives a
XX CC smoother, healthier and youthful looking skin. The method also comprises
XX CC stimulating endogenous production of elastin. The method can be used to
XX CC enhance elasticity of blood vessels or lung tissue. Application of the
XX CC composition results in neovascularisation or angiogenesis.
XX XX Sequence 7 AA;

Query Match 100.0%; Score 27; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5
Db 2 gvgvp 6

RESULT 14
AAY83446
ID AAY83446 standard; peptide; 7 AA.
XX AC AAY83446;
XX DT 29-AUG-2000 (first entry)
XX DE Elastin derived peptide.
XX KW Elastin; peptide; cosmetic; epidermis; skin; dermis;
XX KW dermatological; cutaneous; ageing; steroid; wrinkles; human.
XX

```

```

OS Homo sapiens.
XX PN WO200028996-A1.
XX PD 25-MAY-2000.
XX PF 19-NOV-1999; 99WO-US27399.
XX PR 19-NOV-1998; 98US-0109246.
XX PA (MRSN-) MRS LLC.
XX PI Mitts TF, Sandberg LB, Roos PJ, Dray PJ;
XX XX WPI; 2000-399644/34.
XX PT Cosmetic or dermatological composition comprise 7-hydroxylated steroid
XX PT and optionally elastin peptide fragment, useful for improving tissue
XX PT texture
XX PS Disclosure; Page 48; 74pp; English.
XX CC A cosmetic or dermatological composition comprising a 7-hydroxylated
XX CC steroid and optionally an elastin peptide fragment can be used for
XX CC restoring and preventing cutaneous ageing in persons of a certain
XX CC age as well as providing protective effects in relation to UV,
XX CC firming up the skin and reducing wrinkles. Elastin derived peptides
XX CC are described in GENESEQ records AAY83412-Y83447.
XX XX Sequence 7 AA;

Query Match 100.0%; Score 27; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5
Db 2 gvgvp 6

RESULT 15
AAY06866
ID AAY06866 standard; peptide; 8 AA.
XX AC AAY06866;
XX DT 29-JUN-1999 (first entry)
XX DE Short elastin peptide sequence A.
XX KW Biomolecule; elastomeric; environmental; pharmaceutical; infection;
XX KW extracorporeal treatment; enzyme; polymerase chain reaction; PCR;
XX KW polymerase; fermentation; elastin.
XX OS Synthetic.
XX PN WO9911661-A1.
XX PD 11-MAR-1999.
XX PF 28-AUG-1998; 98WO-GB02602.
XX PR 29-AUG-1997; 97GB-0018463.
XX PA (DYNA-) DYNAL AS.
XX PI Korsnes L, Rees A, Reiersen H;
XX XX WPI; 1999-243602/20.
XX PT Biomolecule containing an inserted elastomeric peptide
XX

```

PS Disclosure; Page 51; 137pp; English.

xx The invention relates to biomolecules comprising elastomeric peptides.
 CC The elastomeric peptide contracts in response to a change in some
 CC environmental parameter. The properties of the functional components in
 CC the biomolecule are altered when contraction or expansion of the
 CC elastomeric peptides is induced. The biomolecules are used: (i) for
 CC purification, based on affinity of its functional component, as an
 CC antibody, and its ability to release bound material after induction of
 CC expansion or contraction; (ii) as a pharmaceutical, for extracorporeal
 CC treatment or localized systemic treatment, e.g. by exploiting variations
 CC in pH in the digestive tract or of temperature at sites of infection;
 CC (iii) as a biosensor to detect changes in the environment of the
 CC molecule, e.g. activation of an enzyme to generate a coloured product,
 CC also in enzyme-linked immunosorbent and other assays, for screening
 CC expression libraries by panning and generally for diagnosis; or (iv)
 CC where the functional component is an enzyme, as reagent for e.g. PCR
 CC (polymerase chain reaction), e.g. a polymerase that becomes active only
 CC at a selected temperature, in fermentations, etc. The elastomeric peptide
 CC provides an inducible switch which may eliminate, increase or reduce
 CC activity; alter activity, e.g. change the kinetic parameters for a
 CC particular enzymatic or binding reaction, or introduce new activity,
 CC especially for control of enzymatic activity, e.g. in PCR as an
 CC alternative to the 'hot start' technique.

xx SQ Sequence 8 AA;

Query Match 100.0%; Score 27; DB 20; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GVGVP 5
 Db 1 GVGVP 5

Search completed: July 25, 2001, 16:49:31
 Job time: 43 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2001, 16:49:56 ; Search time 20.09 Seconds
(without alignments)
5.125 Million cell updates/sec

Title: US-09-251-638-2

Perfect score: 27

Sequence: 1 GVGVP 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	5	1	US-08-106-509-14
2	27	100.0	5	1	US-08-175-155-3
3	27	100.0	5	1	US-08-477-509B-23
4	27	100.0	5	2	US-08-435-641-2
5	27	100.0	5	2	US-08-707-237A-9
6	27	100.0	5	2	US-08-543-020-1
7	27	100.0	5	2	US-08-483-236-17
8	27	100.0	5	2	US-08-735-692-3
9	27	100.0	5	3	US-08-542-051-41
10	27	100.0	5	3	US-08-482-085B-23
11	27	100.0	5	3	US-08-642-246-2
12	27	100.0	5	5	PCT-US96-06229-2
13	27	100.0	5	6	5250516-16
14	27	100.0	7	3	US-09-039-308A-40
15	27	100.0	15	2	US-08-735-692-15
16	27	100.0	16	1	US-08-468-543-17
17	27	100.0	16	2	US-08-469-692-17
18	27	100.0	16	2	US-08-398-046-17
19	27	100.0	20	1	US-08-127-351-35
20	27	100.0	20	1	US-08-480-367B-35
21	27	100.0	20	1	US-08-487-221A-35
22	27	100.0	20	1	US-08-480-370-35
23	27	100.0	20	1	US-08-299-636-14
24	27	100.0	20	1	US-08-279-155-16
25	27	100.0	20	1	US-08-464-456-14
26	27	100.0	20	1	US-08-486-135-2
27	27	100.0	20	1	US-08-703-988A-16

28	27	100.0	20	1	US-08-470-152-2	Sequence 2, Appl
29	27	100.0	20	1	US-08-463-052-14	Sequence 14, Appl
30	27	100.0	20	2	US-08-480-551-14	Sequence 14, Appl
31	27	100.0	20	2	US-08-543-020-12	Sequence 12, Appl
32	27	100.0	20	2	US-08-612-842-16	Sequence 16, Appl
33	27	100.0	20	2	US-08-290-853-23	Sequence 23, Appl
34	27	100.0	20	3	US-08-542-051-21	Sequence 21, Appl
35	27	100.0	21	1	US-08-468-543-13	Sequence 13, Appl
36	27	100.0	21	2	US-08-469-692-13	Sequence 13, Appl
37	27	100.0	21	2	US-08-398-046-13	Sequence 13, Appl
38	27	100.0	22	2	US-08-707-237A-98	Sequence 98, Appl
39	27	100.0	22	3	US-08-642-246-18	Sequence 18, Appl
40	27	100.0	22	5	PCT-US96-06229-18	Sequence 18, Appl
41	27	100.0	23	1	US-08-472-535-13	Sequence 13, Appl
42	27	100.0	23	1	US-08-484-774-13	Sequence 13, Appl
43	27	100.0	23	2	US-08-290-853-8	Sequence 8, Appl
44	27	100.0	23	3	US-08-266-178A-13	Sequence 13, Appl
45	27	100.0	23	6	5250516-5	Patent No. 5250516

ALIGNMENTS

RESULT 1
US-08-106-509-14
; Sequence 14, Application US/08106509
; Patent No. 5428014
; GENERAL INFORMATION:
; APPLICANT: Labroc, Virender
; APPLICANT: Busby, Sharon J.
; TITLE OF INVENTION: TRANSGUTAMINASE CROSS-LINKABLE
; TITLE OF INVENTION: POLYPEPTIDES AND METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/106.509
; FILING DATE: 19930813
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-684
; REFERENCE/DOCKET NUMBER: 93-09
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-106-509-14

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GVGVP 5
Db 1 GVGVP 5

```
RESULT 2
US-08-175-155-3
; Sequence 3, Application US/08175155
; Patent No. 5641648
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph W.
; APPLICANT: Crissman, John W.
; APPLICANT: Dorman, Mary A.
; TITLE OF INVENTION: Methods for Preparing Synthetic
; TITLE OF INVENTION: Repetitive DNA
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/175,155
; FILING DATE: 29-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I.
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-5/BIR
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-175-155-3

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
Db 1 GVGVP 5

RESULT 3
US-08-477-509B-23
; Sequence 23, Application US/08477509B
; Patent No. 5770697
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph W.
; APPLICANT: Crissman, John W.
; APPLICANT: Dorman, Mary A.
; TITLE OF INVENTION: No. 5770697el Peptides Comprising Repetitive
; TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
```

```
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,509B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,155
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,049
FILING DATE: 22-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-7/RFT/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-477-509B-23

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
Db 1 GVGVP 5

RESULT 4
US-08-435-641-2
; Sequence 2, Application US/08435641
; Patent No. 5817303
; GENERAL INFORMATION:
; APPLICANT: Stedronsky, Erwin R.
; TITLE OF INVENTION: Tissue Adhesive Using Synthetic
; TITLE OF INVENTION: Crosslinking
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,641
; FILING DATE: 05-MAY-1995
```

CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Rowland, Bertram I.
 REGISTRATION NUMBER: 20015
 REFERENCE/DOCKET NUMBER: A-61127
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-781-1989
 TELEFAX: 415-398-3249
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-435-641-2

Query Match 100.0%; Score 27; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5
 Db 1 GVGVP 5

RESULT 5
 US-08-707-237A-9
 Sequence 9, Application US/08707237A
 Patent No. 5830713
 GENERAL INFORMATION:
 APPLICANT: Ferrari, Franco A.
 APPLICANT: Capello, Joseph
 APPLICANT: Crissman, John W.
 APPLICANT: Dorman, Mary A.
 TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC
 TITLE OF INVENTION: REPTITIVE DNA
 NUMBER OF SEQUENCES: 108
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
 STREET: Four Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: United States
 ZIP: 94111-4187
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/707,237A
 FILING DATE: 03-SEP-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/175,155
 FILING DATE: 29-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/053,049
 FILING DATE: 22-APR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/609,716
 FILING DATE: 06-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/269,429
 FILING DATE: 03-NOV-1988
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/114,618
 FILING DATE: 29-OCT-1987
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 06/927,258
 FILING DATE: 04-NOV-1986

ATTORNEY/AGENT INFORMATION:
 NAME: Trecartin, Richard F.
 REGISTRATION NUMBER: 31,801
 REFERENCE/DOCKET NUMBER: A-55186-10/WHD
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249
 TELEX: 910 277299
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-707-237A-9

Query Match 100.0%; Score 27; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5
 Db 1 GVGVP 5

RESULT 6
 US-08-543-020-1
 Sequence 1, Application US/08543020
 Patent No. 5854387
 GENERAL INFORMATION:
 APPLICANT: Urry, Dan W.
 APPLICANT: McPherson, David T.
 APPLICANT: Xu, Jie
 TITLE OF INVENTION: A Simple Method for the Purification of
 TITLE OF INVENTION: a Bioelastic Polymer
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
 STREET: 5 Palo Alto Square
 CITY: Palo Alto
 STATE: CA
 COUNTRY: US
 ZIP: 94306-2155
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/543,020
 FILING DATE: 13-OCT-1995
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Hughes, Melya J.
 REGISTRATION NUMBER: 38,696
 REFERENCE/DOCKET NUMBER: BERL-016/010US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415 853 5070
 TELEFAX: 415 857 0663
 TELEX: 380816COOLEYPA
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-543-020-1

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5
Db 1 GVGVP 5

RESULT 7
US-08-483-236-17
; Sequence 17, Application US/08483236
; Patent No. 5939385
; GENERAL INFORMATION:
; APPLICANT: Labroo, Virender
; APPLICANT: Busby, Sharon
; TITLE OF INVENTION: Transglutaminase Cross-Linkable
; TITLE OF INVENTION: Polypeptides and Methods Relating Thereto
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,236
; FILING DATE:
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-09c1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: Single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-483-236-17

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5
Db 1 GVGVP 5

RESULT 8
US-08-735-692-3
; Sequence 3, Application US/08735692B
; Patent No. 5972406
; GENERAL INFORMATION:
; APPLICANT: Urry, Dan W.
; APPLICANT: Shewry, Peter R.
; APPLICANT: Prasad, Kari
; TITLE OF INVENTION: Bielastromers Suitable as Food Product Additives
; FILE REFERENCE: BERL-017/01US
; CURRENT APPLICATION NUMBER: US/08/735,692B
; CURRENT FILING DATE: 1995-10-16

; EARLIER APPLICATION NUMBER: 08/423,517
; EARLIER FILING DATE: 1995-04-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
US-08-735-692-3

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5
Db 1 GVGVP 5

RESULT 9
US-08-542-051-41
; Sequence 41, Application US/08542051F
; Patent No. 6004782
; GENERAL INFORMATION:
; APPLICANT: Daniell, Henry
; APPLICANT: McPherson, David T.
; APPLICANT: Urry, Dan W.
; APPLICANT: Xu, Jie
; TITLE OF INVENTION: Hyperexpression of Bioelastic Polypeptides
; FILE REFERENCE: BERL-018/01US
; CURRENT APPLICATION NUMBER: US/08/542,051F
; CURRENT FILING DATE: 1995-10-13
; EARLIER APPLICATION NUMBER: 08/423,642
; EARLIER FILING DATE: 1995-04-14
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
US-08-542-051-41

Query Match 100.0%; Score 27; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVGVP 5
Db 1 GVGVP 5

RESULT 10
US-08-482-085B-23
; Sequence 23, Application US/08482085B
; Patent No. 6018030
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Richardson, Charles
; APPLICANT: Chambers, James
; APPLICANT: Causey, Stuart
; APPLICANT: Pollock, Thomas J.
; APPLICANT: Cappelletto, Joseph
; APPLICANT: Crissman, John W.
; TITLE OF INVENTION: NO. 6018030a1 Peptides Comprising Repetitive
; TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/482,085B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,049
; FILING DATE: 22-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,155
; FILING DATE: 29-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Irecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-6/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-482-085B-23

```

```

Query Match 100.0%; Score 27; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GVGVP 5
   11111
DB 1 GVGVP 5

```

```

RESULT 11
US-08-642-246-2
; Sequence 2, Application US/08642246
; Patent No. 6033654
; GENERAL INFORMATION:
; APPLICANT: STEDRONSKY, Erwin R.
; APPLICANT: CAPPELLO, Joseph
; TITLE OF INVENTION: Tissue Adhesive Using Synthetic
; TITLE OF INVENTION: Crosslinking
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: Four Embarcadero Center, Suite 200
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,246
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A61127-1/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-642-246-2

```

```

Query Match 100.0%; Score 27; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GVGVP 5
   11111
DB 1 GVGVP 5

```

```

RESULT 12
PCT-US96-06229-2
; Sequence 2, Application PC/TUS9606229
; GENERAL INFORMATION:
; APPLICANT: STEDRONSKY, Erwin R.
; APPLICANT: CAPPELLO, Joseph
; TITLE OF INVENTION: Tissue Adhesive Using Synthetic
; TITLE OF INVENTION: Crosslinking
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: Four Embarcadero Center, Suite 200
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/06229
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A61127-1/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US96-06229-2

```

Query Match 100.0%; Score 27; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
Db 1 GVGVP 5

RESULT 13
5250516-16
; Patent No. 5250516
; APPLICANT: URRY, DAN W.
; TITLE OF INVENTION: BIOLASTOMERIC MATERIALS SUITABLE FOR
; THE PROTECTION OF BURN AREAS OR THE PROTECTION OF WOUND
; REPAIR SITES FROM THE OCCURRENCE OF ADHESIONS
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/184,407
; FILING DATE: 21-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 900,895
; FILING DATE: 27-AUG-1986
; APPLICATION NUMBER: 853,212
; FILING DATE: 17-APR-1986
; SEQ ID NO:16:
; LENGTH: 5
5250516-16

Query Match 100.0%; Score 27; DB 6; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
Db 1 GVGVP 5

RESULT 14
US-09-039-308A-40
; Sequence 40, Application US/09039308A
; Patent No. 6069129
; GENERAL INFORMATION:
; APPLICANT: Sandberg, Lawrence; Roos, Phillip;
; TITLE OF INVENTION: Mitts, Thomas
; TITLE OF INVENTION: ELASTIN DERIVED COMPOSITION
; TITLE OF INVENTION: AND METHOD OF
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED SMITH SHAW & MCCLAY, LLP
; STREET: PO Box 488
; CITY: Pittsburgh
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 15230
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
; COMPUTER: Compaq
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,308A
; FILING DATE: March 13, 1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Miller, Raymond A.
; REGISTRATION NUMBER: 42,891
; REFERENCE/DOCKET NUMBER: 97-489
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (412) 288-4192

TELEFAX: (412) 288-3300
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-039-308A-40

Query Match 100.0%; Score 27; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
Db 2 GVGVP 6

RESULT 15
US-08-735-692-15
; Sequence 15, Application US/08735692B
; Patent No. 5972406
; GENERAL INFORMATION:
; APPLICANT: Urry, Dan W.
; APPLICANT: Shewry, Peter R.
; APPLICANT: Prasad, Karl
; TITLE OF INVENTION: Bielasomers Suitable as Food Product Additives
; FILE REFERENCE: BERL-017/01US
; CURRENT APPLICATION NUMBER: US/08/735,692B
; CURRENT FILING DATE: 1995-10-16
; EARLIER APPLICATION NUMBER: 08/423,517
; EARLIER FILING DATE: 1995-04-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
US-08-735-692-15

Query Match 100.0%; Score 27; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVGVP 5
Db 11 GVGVP 15

Search completed: July 25, 2001, 16:49:57
Job time: 68 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2001, 16:48:48 ; Search time 32.96 Seconds
(without alignments)
9.197 Million cell updates/sec

Title: US-09-251-638-1

Perfect score: 27

Sequence: 1 VPGVG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0601.*

```

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	27	100.0	5	7 AAP61199	Repeating unit.
2	27	100.0	5	13 AAR29149	Pentapeptide repea
3	27	100.0	5	13 AAR29145	Bioelastic pentape
4	27	100.0	5	16 AAR5228	Fibrinogen/beta-ca
5	27	100.0	5	16 AAR80250	Elastin-like repet
6	27	100.0	5	16 AAR80307	Elastin cross-link
7	27	100.0	5	18 AAW22714	Bioelastomeric rep
8	27	100.0	5	18 AAW26332	Elastin repeat uni
9	27	100.0	5	18 AAW18261	Transglutaminase c
10	27	100.0	5	18 AAW12301	Monomeric unit for
11	27	100.0	5	19 AAW49701	Elastin repeat mot

12	27	100.0	5	19 AAW53506	Peptide reapeat un
13	27	100.0	5	19 AAW47348	Elastin repeat uni
14	27	100.0	5	20 AAV55877	Apolipoprotein fra
15	27	100.0	5	20 AAY31886	Pentapeptide used
16	27	100.0	5	20 AAY33102	Elastin derived re
17	27	100.0	5	21 AAB19211	Elastin fragment u
18	27	100.0	5	21 AAY80334	Repeat elastin pen
19	27	100.0	5	21 AAY78246	Repeating unit SEQ
20	27	100.0	5	21 AAV78265	Antigen peptide se
21	27	100.0	5	22 AAB72702	Repetitive protein
22	27	100.0	5	22 AAB63972	Elastin repeating
23	27	100.0	6	19 AAW47335	Inhibitor of LDL b
24	27	100.0	6	20 AAY55874	Apolipoprotein fra
25	27	100.0	6	20 AAY33099	Elastin derived pe
26	27	100.0	8	20 AAY06866	Short elastin pept
27	27	100.0	9	20 AAY06867	Short elastin pept
28	27	100.0	13	10 AAP91307	Sequence of beta-t
29	27	100.0	13	18 AAW34883	Bioelastomeric pol
30	27	100.0	15	20 AAY55862	Apolipoprotein fra
31	27	100.0	16	19 AAW47350	Inhibitor of LDL b
32	27	100.0	16	20 AAY55879	Apolipoprotein fra
33	27	100.0	16	20 AAY33104	Elastin derived re
34	27	100.0	18	19 AAW44502	Bovine "polyprollin
35	27	100.0	18	20 AAY06871	Short elastin pept
36	27	100.0	19	20 AAY06872	Short elastin pept
37	27	100.0	20	13 AAR26383	Sequence of peptid
38	27	100.0	20	14 AAR40000	Scintigraph imagin
39	27	100.0	20	17 AAW11073	Atherosclerotic pl
40	27	100.0	20	18 AAW31123	Atherosclerotic pl
41	27	100.0	20	20 AAY55861	Apolipoprotein fra
42	27	100.0	20	20 AAY06873	Short elastin pept
43	27	100.0	21	12 AAR15136	Vascular injury af
44	27	100.0	21	14 AAR42543	Leukocyte-binding
45	27	100.0	21	19 AAW47336	Inhibitor of LDL b

ALIGNMENTS

```

RESULT 1
AAP61199
ID AAP61199 standard; protein; 5 AA.
AC AAP61199;
XX
XX
DT 01-AUG-1991 (first entry)
XX
DE Repeating unit.
XX
KW Elastomeric copolymer; prosthesis; repeating unit.
XX
PN US4589882-A.
XX
PD 20-MAY-1986.
XX
PF 19-SEP-1983; 83US-0533524.
XX
PR 19-SEP-1983; 83US-0533524.
XX
PA (URRY/) URRY D W.
XX
PI Urry DW;
XX
DR WPI; 1986-149845/23.
XX
PT Synthetic elastomeric copolymers - useful as prostheses for
PT repair of ligaments, tendons and blood vessel walls
XX
PS Claim 22; page 10; lipp; English.
XX
CC The repeating unit can be used in the prodn. of a synthetic elastic
CC copolymer. The copolymer contains a beta turn and the repeating
CC unit contains hydrophobic amino acid and glycine residues and a cross-

```

CC linking component. The copolymer is useful in prosthetic systems, for
 CC repairing a natural elastic system. It is so functionalised so as to
 CC provide reactive gps. which can become covalently cross-linked by
 CC tissue enzymes to newly synthesised connective tissue protein.

XX Sequence 5 AA;
 SQ Query Match 100.0%; Score 27; DB 7; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 |||||
 Db 1 vpgvg 5

RESULT 2
 AAR29149
 ID AAR29149 standard; peptide; 5 AA.
 XX AC AAR29149;
 XX DT 06-MAY-1993 (first entry)
 XX Pentapeptide repeating unit of bioelastic polymer.
 XX Superabsorbent; bioelastic; diaper; hygienic articles; wound;
 KW dressing; implant; inverse temperature transition; tissue; napkin;
 KW catbox liner; toilet paper; toweltete; cleaning wipe; bandage;
 KW medical sponge; swab; printing ink; contact lens.

XX Synthetic.
 XX WO9218079-A.
 XX 29-OCT-1992.
 XX 10-MAR-1992; 92WO-US01959.
 XX 19-APR-1991; 91US-0688185.
 XX (BIOE-) BIOELASTICS RES LTD.
 XX Urry DW;
 XX WPI; 1992-381725/46.
 XX Super-absorbent material incorporating polymer undergoing inverse
 PT temp. transition - esp. bio-elastic polypeptide(s) for
 PT controllably absorbing body fluids

XX Claim 27/28; Page 34; 42pp; English.
 XX The invention relates to an appliance suitable for contacting body
 CC surfaces and for absorbing aqueous liquids including body fluids.
 CC The appliance includes a polymeric material which undergoes an
 CC inverse temperature transition, pref. such that it is in a
 CC contracted state at a higher temperature and in a swollen state when
 CC at a lower temperature. The appliance, e.g. a diaper, may be
 CC selected to be in a contracted state prior to use and while
 CC contacting the body surface, and to be in a swollen state after
 CC absorbing a body fluid and being at a new location having a lower
 CC temperature distant from the body-contacting location. Pref. the
 CC polymeric material is a bioelastic polymer containing elastomeric
 CC tetrapeptide, pentapeptide or nonapeptide repeat units, the polymer
 CC having a series of beta turns separated by dynamic bridging segments
 CC suspended between the beta turns. In particular, the polymer comprises a
 CC segment of formula poly[(VPGXG)x(VPGVG)y], in which x and y are mole
 CC fractions such that x+y = 1, and x is a hydrophobic amino acid residue;
 CC or of formula poly[(VPGXG)x(VPGVG)y(VPGZG)z], in which x+y+z = 1, x = as
 CC above, and z is an amino acid residue having a side chain capable of
 CC undergoing reversible protonation in an aqueous environment.

CC A specific polymer is poly(VPGVG).
 XX Sequence 5 AA;
 SQ

Query Match 100.0%; Score 27; DB 13; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 |||||
 Db 1 vpgvg 5

RESULT 3
 AAR29145
 ID AAR29145 standard; peptide; 5 AA.
 XX AC AAR29145;
 XX DT 23-MAR-1993 (first entry)
 XX Bioelastic pentapeptide polymer fragment 2.
 KW Polymer fragment; inverse temperature transition; exerted pressure;
 KW mechanical work; chemical change; hydrophobicity.

XX Synthetic.
 XX WO9219183-A.
 XX 12-NOV-1992.
 XX 03-APR-1992; 92WO-US02691.
 XX 22-APR-1991; 91US-0688324.
 XX (UABR-) UAB RES FOUND.
 XX Urry DW;
 XX WPI; 1992-398469/48.

XX Pressure expanding polymers giving reversible mechanical changes
 PT - has inverse temp. transition in liq. water range and pref.
 PT being bio-elastic polypeptide(s) contg. hydrophobic gps.
 XX Claims 13 and 14; Page 36; 51pp; English.

XX The sequences given in AAR29144-46 are polymer fragments that are used
 CC in polymeric materials which have an inverse temp. transition in the
 CC range of liq. water and which expand against an exerted pressure.
 CC Fragments 1 and 2 are used such that mole fractions equal 1 in one
 CC polymeric material and fragments 1, 2 and 3 are used in a further
 CC polymeric material such that mole fractions equal 1. The materials
 CC can be used in a variety of applications to produce mechanical work
 CC and/or cause chemical changes in a sealed environment by variation of
 CC the pressure on the material. The degree of mechanical or chemical
 CC change can be controlled by selection of the number, hydrophobicity
 CC and size of the hydrophobic groups and the presence or absence of
 CC reactive functional groups in the polymer.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 13; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 |||||
 Db 1 vpgvg 5

XX Cappello J;
 XX WPI; 1995-320413/41.
 XX Protein polymers comprising repeating units and sequences - capable
 PT of enzyme-catalysed covalent bond formation useful as a
 PT biocompatible material for wound closure and tissue repair
 XX
 XX Disclosure; Page 12; 138pp; English.
 XX
 CC The sequence of the cross-linking reactive motif from elastin. The motif
 CC can be used in a novel polymer comprising two spaced enzyme recognition
 CC site and may contain repetitive units of 3-8 amino acids with at least
 CC two pendant groups. The polymers contg. the multimeric repeat sequence
 CC can be used as substrates for enzymatic cross-linking. The polymers can
 CC be used in biological systems where in situ formation of a biocompatible
 CC material with structural integrity is required e.g. as medical adhesives
 CC and sealants or for wound closure or tissue repair.
 XX
 XX Sequence 5 AA;
 SQ

Query Match 100.0%; Score 27; DB 16; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 DB 1 vpgvg 5

RESULT 7
 AAW22714
 ID AAW22714 standard; peptide; 5 AA.
 XX
 AC AAW22714;
 DT 26-FEB-1998 (first entry)
 DE Bioelastomeric repeating unit 1 responsive to electrical energy.
 XX
 KW Bioelastomeric repeating unit; polymer; inverse temperature transition;
 KW electrical energy change; polarity; hydrophobicity; mechanical work;
 KW surgical suture; heat shrinkable membrane; desalination.
 XX
 OS Synthetic.
 XX
 PN WO9723729-A1.
 PD 03-JUL-1997.
 XX
 PF 07-JUN-1996; 96WO-US09776.
 XX
 PR 07-JUN-1995; 95US-0487594.
 XX
 PA (URRY/) URRY D W.
 XX
 PI Urry DW;
 XX
 DR WPI; 1997-363360/33.
 XX
 PT Bioelastic polymer responsive to electrical energy - comprising
 PT beta turn and residue(s) with side chain that changes polarity or
 PT hydrophobicity in response to electrical energy change, useful for
 PT mechanical work or light stimulated contraction
 XX
 PS Claim 13; Page 52; 60pp; English.
 XX
 CC AAW22711-16 represent bioelastomeric repeating units that are used to
 CC create a polymer responsive to electrical energy. This bioelastic
 CC polymer comprises a bioelastomeric polypeptide having an inverse
 CC temperature transition, and a repeating bioelastomeric unit containing

CC at least 1 beta-turn and residues with a side chain. The side chain of
 CC these residues responds to an electrical energy change by altering its
 CC polarity or hydrophobicity. The side chain is present in sufficient
 CC amounts to provide a shift in the inverse temperature transition of the
 CC polymer upon the change in exposure to electrical energy. The polymer may
 CC also contain a second amino acid with a side chain capable of undergoing
 CC a change in an aqueous environment. The polymer (AAW34881-83) consists of
 CC the formulas described in AAW22716 and AAW22711 and below:

CC poly[fx(VPGXG),fv(VPDVG)] where

CC fx and fv are mole fractions with fx + fv = 1

CC x represents an amino acid residue having an electrically responsive
 CC side chain.

CC The expansion or contraction of the polymer upon exposure to a change
 CC in electrical energy can be used to produce mechanical work when the
 CC polymer is constrained. It can also cause turbidity and chemical changes
 CC in a closed environment. The polymer can be used in surgical sutures,
 CC especially for microsurgery, heat shrinkable membranes, controlled
 CC herbicide or pesticide release or for desalination.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 18; Length 5;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

DB 1 vpgvg 5

RESULT 8

AAW26332

ID AAW26332 standard; Peptide; 5 AA.

XX

AC AAW26332;

DT 19-NOV-1997 (first entry)

DE Elastin repeat unit.

XX Elastin; protein polymer.

XX Synthetic.

OS

PN US5641648-A.

XX

PD 24-JUN-1997.

XX

PF 04-NOV-1986; 86US-0927258.

XX

PR 29-DEC-1993; 93US-0175155.

XX

PR 04-NOV-1986; 86US-0927258.

XX

PR 29-OCT-1987; 87US-0114618.

XX

PR 06-NOV-1988; 88US-0269429.

XX

PR 06-NOV-1990; 90US-0609716.

XX

PR 22-APR-1993; 93US-0053049.

XX

PA (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

XX

PI Cappello J, Ferrari FA, Richardson C;

XX

DR WPI; 1997-340943/31.

XX

PT Preparation of synthetic DNA encoding a protein comprising short

XX repeats - by synthesising oligomers, annealing and oligomerising

PT these, particularly to produce proteins that mimic silk, collagen

XX etc

PS Claim 8; Column 11; 90pp; English.

XX

CC This peptide represents a repeat unit found in elastin. Methods

CC are claimed for preparing protein polymers that contain repeating

CC units based on those found in elastin, fibroin, keratin, etc. (see
 CC also AAW26329, AAW26334 and AAW26336). The method involves:
 CC synthesizing pairs of single-stranded oligomers, each member of the pair
 CC overlapping except at the protruding ends; hybridising each pair to
 CC double-stranded segments; combining the double-stranded segments, or
 CC their cloned copies, in a cloning vector to form a monomer; excising
 CC the monomer from the vector; and oligomerising the monomer to
 CC produce a multimer. Such polymers comprise fibrous or structural
 CC proteins, including crystalline, elastomeric, tough and bony
 CC materials such as those that mimic (and can substitute for) silk,
 CC elastin, collagen, keratin etc. The properties of the protein can
 CC be controlled by varying the type of units in the monomer, the
 CC number of units per multimer, the spacing between them and the
 CC number of multimer repeats.
 XX
 XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 18; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 DB 1 VPGVG 5

RESULT 9
 AAW18261
 ID AAW18261 standard; peptide; 5 AA.

XX AC AAW18261;
 XX DT 29-AUG-1997 (first entry)

DE Transglutaminase cross-linkable polypeptide elastomeric peptide.
 XX
 XX Elastomeric; homopolymer; copolymer; tissue sealant; skin graft;
 KW burn.

XX Synthetic.
 OS WO9640780-A1.
 PN 19-DEC-1996.

XX PF 31-MAY-1996; 96WO-US08269.
 XX PR 07-JUN-1995; 95US-0483236.

XX PA (ZYMO) ZYMOGENETICS INC.
 XX PI Busby SJ, Labroo VM;
 XX WPI; 1997-052237/05.

PT Transglutaminase cross-linkable peptide(s) - used in the mfr. of
 PT biocompatible, bio:adhesive tissue sealant and wound healing
 PT preparations.

PS Claim 6; Page 48; 55pp; English.

XX Novel polypeptides optionally having one or both of the amino-terminus
 CC and carboxy-terminus flanked by an elastomeric peptide, which is cross-
 CC linkable by a transglutaminase comprises a segment of formula:
 CC SI-Y-S2. The present sequence represents a specifically claimed
 CC example of an elastomeric peptide. The homo- and copolymers produced
 CC are useful in tissue sealant and wound healing formulations. Tissue
 CC sealants are useful in skin grafting for burn victims and for sealing
 CC surgical and other wounds.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 18; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 DB 1 VPGVG 5

RESULT 10
 AAW12301
 ID AAW12301 standard; peptide; 5 AA.

XX AC AAW12301;
 XX DT 22-APR-1997 (first entry)

XX Monomeric unit for elastic protein-base polymer for chewing gum..
 DE Chewing gum; water; insoluble; plasticiser; elastomer; flavour; polymer;
 KW sweetener; texturiser; crosslinkage; gamma-irradiation; drug; nutrient;
 KW environmentally friendly; drug; anti-inflammatory agent; vitamin.

XX Synthetic.
 OS US5580590-A.

XX PN 03-DEC-1996.
 XX PD 27-DEC-1993; 93US-0174185.

XX PR 27-DEC-1993; 93US-0174185.
 XX PA (WRIL) WRIGLEY JR CO WM.

XX PI Hartman SE;
 XX WPI; 1997-033531/03.

PT Environmentally friendly chewing gum - contg. water insoluble
 PT elastic polypeptide having a penta:peptide repeat

XX Claim 1; Column 7; 6pp; English.

XX A novel chewing gum composition comprises a water insoluble polymer
 CC comprising this pentapeptide as a base. The insoluble portion can
 CC comprise 1-99% of the gum, together with a plasticiser, an elastomer,
 CC a flavour, a sweetener and a texturiser. The peptide is preferably
 CC crosslinked by gamma-irradiation to form the polymer. The new chewing
 CC gum is environmentally friendly as it can be swallowed after chewing or
 CC easily removed from surfaces. The peptide can also have drugs or other
 CC nutrients e.g. anti-inflammatory agents, vitamins, etc., attached to the
 CC peptide, for therapeutic purposes.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 18; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 DB 1 VPGVG 5

RESULT 11
 AAW49701
 ID AAW49701 standard; Peptide; 5 AA.

XX AC AAW49701;
 XX

DT 12-OCT-1998 (first entry)
 XX
 DE Elastin repeat motif.
 XX
 KW Protein polymer; cross-linking; elastin; adhesive; sealant;
 KW wound healing.
 XX
 OS Synthetic.
 XX
 PN US5773577-A.
 XX
 PD 30-JUN-1998.
 XX
 PF 03-MAR-1994; 94US-0205518.
 XX
 PR 02-MAR-1995; 95US-0397633.
 PR 03-MAR-1994; 94US-0205518.
 XX
 PA (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
 XX
 PI Cappello J;
 XX
 DR WPI; 1998-387091/33.
 XX

XX New recombinant protein polymers - containing naturally occurring
 PT repetitive units for crosslinking by enzymes, useful as medical
 PT adhesives and sealants, depots and matrices
 XX
 PS Disclosure; Column 7; 70pp; English.
 XX
 CC This is a repeat motif of elastin. A claimed recombinant protein
 CC polymer of 15-250 kDa comprises a repetitive amino acid backbone of
 CC repetitive units having a collagen, fibroin, elastin or keratin
 CC motif and at least 2 enzyme recognition sequences comprising a
 CC glutamine capable of enzyme-catalysed isopeptide formation,
 CC separated by an intervening sequence of at least 25 amino acids.
 CC Such recombinant protein polymers are capable of covalent
 CC crosslinking by enzymatic reaction to form products which set
 CC quickly and have good adhesive properties and high strength. The
 CC compositions can be used as medical adhesives and sealants, in the
 CC closure of wounds and repair of damaged tissues, prostheses
 CC coatings, drug depots, and matrices for the transplantation of
 CC cells. They can also be used in assays for analytes.
 XX
 SQ Sequence 5 AA;

Query Match 100.0%; Score 27; DB 19; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 Db 1 VPGVG 5

RESULT 12
 AAW53506
 ID AAW53506 standard; Protein; 5 AA.
 XX
 AC AAW53506;
 XX
 DT 10-AUG-1998 (first entry)
 XX
 DE Peptide repeat unit 4.
 XX
 KW Peptide repeat unit; DNA repeat unit; high molecular weight polymer;
 KW synthetic silk; silk worm; elastin.
 XX
 OS Synthetic.
 OS Bombyx mori.
 XX
 PN WO9810063-A1.

XX
 PD 12-MAR-1998.
 XX
 PF 23-SEP-1996; 96WO-US15306.
 XX
 PR 03-SEP-1996; 96US-0707237.
 XX
 PA (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
 XX
 PI Cappello J, Crissman JW, Dorman MA, Ferrari FA;
 XX
 DR WPI; 1998-193613/17.
 XX

XX Preparation of synthetic repetitive DNA - useful for construction of
 PT large protein polymers having repeating units, used in structural
 PT material; e.g. synthetic silk
 XX
 PS Claim 11; Page 22; 127pp; English.
 XX
 CC This is the amino acid sequence of the peptide repeat unit, which
 CC is generally found in elastin, and used in the method of the invention,
 CC which involves the preparation of synthetic DNA sequence having of at
 CC repeating units from about 3-15 codons and encoding a protein of at
 CC least about 30 kDa. The method is useful for the production of high
 CC molecular weight polymers (e.g. synthetic silk), either nucleic acids
 CC or peptides that are the expression products of the nucleic acids and
 CC particularly high molecular weight peptides containing repeating units
 CC which are useful as structural materials.
 XX
 SQ Sequence 5 AA;

Query Match 100.0%; Score 27; DB 19; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 Db 1 VPGVG 5

RESULT 13
 AAW47348
 ID AAW47348 standard; peptide; 5 AA.
 XX
 AC AAW47348;
 XX
 DT 01-JUN-1998 (first entry)
 XX
 DE Elastin repeat unit.
 XX
 KW Binding inhibitor; low-density lipoprotein; LDL; vascular wall;
 KW vascular injury; elastin; collagen; prevention; treatment;
 KW vascular disease; atherosclerosis; repeat unit.
 XX
 OS Homo sapiens.
 OS
 PN US5726153-A.
 XX
 PD 10-MAR-1998.
 XX
 PF 06-JUN-1995; 95US-0468543.
 XX
 PR 06-JUN-1995; 95US-0468543.
 PR 02-MAY-1988; 88US-0189130.
 PR 03-MAY-1990; 90US-0518142.
 PR 03-MAY-1990; 90US-0518215.
 PR 02-MAY-1991; 91US-0694929.
 PR 16-APR-1993; 93US-0048569.
 PR 24-FEB-1994; 94US-0201057.
 PR 28-FEB-1995; 95US-0398046.
 PR 06-JUN-1995; 95US-0468543.
 XX

PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Findeis MA, Fischman A, Lees AM, Lees RS, Shih I;
 XX WPI; 1998-192802/17.
 DR Inhibiting binding of low-density lipoprotein to vascular walls - by
 PT administering peptide with affinity for vascular injury sites
 XX Disclosure; Column 14; 31pp; English.
 XX A novel method for inhibiting the binding of a low-density
 CC lipoprotein (LDL) to vascular walls in vivo, comprises
 CC administering a synthetic water soluble peptide containing an
 CC amphiphilic domain and having affinity for sites of vascular
 CC injury, e.g. derived from the present peptide.
 CC The peptide inhibits the binding of LDL to vascular wall
 CC components, e.g. elastin and collagen, and so can be used to
 CC prevent or treat vascular diseases, e.g. atherosclerosis.
 XX Sequence 5 AA;
 PS Query Match 100.0%; Score 27; DB 19; Length 5;
 XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX QY 1 VPGVG 5
 XX Db 1 VPGVG 5
 XX RESULT 14
 XX AAY55877
 ID AAY55877 standard; peptide; 5 AA.
 XX AC AAY55877;
 XX 01-FEB-2000 (first entry)
 XX DE Apolipoprotein fragment peptide #27 for vascular disease imaging.
 XX Diagnosis; water soluble; amphiphilic domain; affinity; vascular injury;
 KW detection; imaging; disease; atherosclerosis; apolipoprotein.
 XX Synthetic.
 OS Homo sapiens.
 XX US5972890-A.
 XX 26-OCT-1999.
 XX PF 28-FEB-1995; 95US-0398046.
 XX 02-MAY-1991; 91US-0694929.
 PR 16-APR-1993; 93US-0048569.
 PR 24-FEB-1994; 94US-0201057.
 PR 02-MAY-1988; 88US-0189130.
 PR 03-MAY-1990; 90US-0518142.
 PR 03-MAY-1990; 90US-0518215.
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Lees AM, Fischman A, Shih I, Findeis MA, Lees RS;
 XX WPI; 1999-632641/54.
 XX New diagnostic synthetic peptides which have affinity for and
 PT accumulate at a site of vascular injury useful for detection and
 PT imaging of vascular disease such as atherosclerosis -
 XX Claim 37; Column 34; 30pp; English.
 XX

CC The peptides AAY55851-Y55889 represent examples of diagnostic, synthetic
 CC peptides which carry a detectable label, contain 30 or fewer amino
 CC acids, are water soluble, contain an amphiphilic domain and have affinity
 CC for, and propensity to accumulate at, a site of vascular injury. They
 CC are preferably derived from the amino acid sequence of apolipoprotein.
 CC The peptides can be used for the detection or imaging of a vascular
 CC injury or disease, e.g. atherosclerosis.
 XX Sequence 5 AA;
 PS Query Match 100.0%; Score 27; DB 20; Length 5;
 XX Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX QY 1 VPGVG 5
 XX Db 1 VPGVG 5
 XX RESULT 15
 XX AAY31686
 ID AAY31686 standard; Peptide; 5 AA.
 XX AC AAY31686;
 XX 22-NOV-1999 (first entry)
 XX DE Pentapeptide used in novel elastomer polymers.
 XX Elastomer; bioelastomer; polymer; tissue augmentation;
 KW tissue restoration; tissue reconstruction; tissue repair; implant.
 XX Synthetic.
 OS WO9943271-A1.
 XX 02-SEP-1999.
 XX PF 26-FEB-1999; 99WO-US04440.
 PR 29-MAY-1998; 98US-0087155.
 PR 27-FEB-1998; 98US-0076297,
 XX (BIOE-) BIOELASTICS RES LTD.
 XX Glazer PA, Parker TM, Urry DW;
 XX WPI; 1999-540487/45.
 XX Augmentation or restoration of mammalian tissue by injecting
 PT solution of peptide polymer, used for soft or hard tissue
 PT reconstruction, especially of intervertebral disks
 XX Claim 9; Page 75; 133pp; English.
 XX The present sequence represents an example of a pentapeptide that
 CC is used in novel bioelastic polymers. The invention provides a
 CC method of tissue augmentation by injecting a polymer comprising
 CC repeating peptide monomeric units selected from nonapeptide,
 CC pentapeptide and tetrapeptide monomeric units, where the monomeric
 CC units form a series of beta-turns separated by dynamic bridging
 CC segments. The polymer has an inverse temperature transition value
 CC that is less than the tissue temp. and is injected in water
 CC solution at coacervate concentration. The polymer can be injected
 CC at peritumoral or subdermal sites (for treatment of urinary
 CC incontinence or for cosmetic purposes), or into hard or soft
 CC tissue, e.g. for repair of traumatic injury. A specific
 CC application is restoration of intervertebral discs.
 XX Sequence 5 AA;
 PS

Query Match 100.0%; Score 27; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
Db 1 vpgvg 5

Search completed: July 25, 2001, 16:49:30
Job time: 42 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2001, 16:48:49 ; Search time 20.09 Seconds
(without alignments)
5.125 Million cell updates/sec

Title: US-09-251-638-1

Perfect score: 27

Sequence: 1 VPGVG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	5	1 US-08-106-509-5	Sequence 5, Appl
2	27	100.0	5	1 US-07-609-716-2	Sequence 2, Appl
3	27	100.0	5	1 US-08-174-185-1	Sequence 1, Appl
4	27	100.0	5	1 US-08-212-237-2	Sequence 2, Appl
5	27	100.0	5	1 US-08-175-155-4	Sequence 4, Appl
6	27	100.0	5	1 US-08-175-155-9	Sequence 9, Appl
7	27	100.0	5	1 US-08-468-543-15	Sequence 15, Appl
8	27	100.0	5	1 US-08-477-509B-24	Sequence 24, Appl
9	27	100.0	5	1 US-08-397-633A-1	Sequence 1, Appl
10	27	100.0	5	1 US-08-397-633A-8	Sequence 8, Appl
11	27	100.0	5	2 US-08-707-237A-10	Sequence 10, Appl
12	27	100.0	5	2 US-08-483-236-15	Sequence 15, Appl
13	27	100.0	5	2 US-08-469-692-15	Sequence 15, Appl
14	27	100.0	5	2 US-08-911-364-6	Sequence 6, Appl
15	27	100.0	5	2 US-08-735-692-20	Sequence 20, Appl
16	27	100.0	5	2 US-08-398-046-15	Sequence 15, Appl
17	27	100.0	5	3 US-08-542-051-1	Sequence 1, Appl
18	27	100.0	5	3 US-08-482-085B-24	Sequence 24, Appl
19	27	100.0	5	3 US-08-482-085B-43	Sequence 43, Appl
20	27	100.0	5	3 US-08-963-168C-27	Sequence 27, Appl
21	27	100.0	5	4 US-08-475-411A-2	Sequence 2, Appl
22	27	100.0	5	4 US-08-478-029A-2	Sequence 2, Appl
23	27	100.0	5	5 PCT-US95-02772-2	Sequence 2, Appl
24	27	100.0	5	5 PCT-US95-02772-2	Sequence 2, Appl
25	27	100.0	5	6 5250516-1	Patent No. 5250516
26	27	100.0	5	6 5250516-17	Patent No. 5250516
27	27	100.0	6	1 US-08-468-543-12	Sequence 12, Appl
27	27	100.0	6	2 US-08-469-692-12	Sequence 12, Appl

28 27 100.0 6 2 US-08-398-046-12 Sequence 12, Appl

29 27 100.0 11 6 5250516-6 Patent No. 5250516

30 27 100.0 16 1 US-08-468-543-17 Sequence 17, Appl

31 27 100.0 16 2 US-08-469-692-17 Sequence 17, Appl

32 27 100.0 16 2 US-08-398-046-17 Sequence 17, Appl

33 27 100.0 20 1 US-08-127-351-35 Sequence 35, Appl

34 27 100.0 20 1 US-08-480-367B-35 Sequence 35, Appl

35 27 100.0 20 1 US-08-487-221A-35 Sequence 35, Appl

36 27 100.0 20 1 US-08-480-370-35 Sequence 35, Appl

37 27 100.0 20 1 US-08-299-636-14 Sequence 14, Appl

38 27 100.0 20 1 US-08-279-155-16 Sequence 16, Appl

39 27 100.0 20 1 US-08-464-456-14 Sequence 14, Appl

40 27 100.0 20 1 US-08-464-135-2 Sequence 2, Appl

41 27 100.0 20 1 US-08-703-988A-16 Sequence 16, Appl

42 27 100.0 20 1 US-08-470-152-2 Sequence 2, Appl

43 27 100.0 20 1 US-08-463-052-14 Sequence 14, Appl

44 27 100.0 20 2 US-08-480-551-14 Sequence 14, Appl

45 27 100.0 20 2 US-08-612-842-16 Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-08-106-509-5

; Sequence 5, Application US/08106509

; Patent No. 5428014

; GENERAL INFORMATION:

; APPLICANT: Labroo, Virender

; APPLICANT: Busby, Sharon J.

; TITLE OF INVENTION: TRANSGUTAMINASE CROSS-LINKABLE

; TITLE OF INVENTION: POLYPEPTIDES AND METHODS RELATING THERETO

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ZymoGenetics, Inc.

; STREET: 4225 Roosevelt Way, N.E.

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98105

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/106,509

; FILING DATE: 19930813

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Parker, Gary E

; REGISTRATION NUMBER: 31-684

; REFERENCE/DOCKET NUMBER: 93-09

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-547-8080 ext 322

; TELEFAX: 206-548-2329

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

US-08-106-509-5

Query Match 100.0%; Score 27; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.5e-05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5

Db 1 VPGVG 5

RESULT 2
US-07-609-716-2
; Sequence 2, Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-3/BIR
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-609-716-2

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
| | | | |
DB 1 VPGVG 5

RESULT 3
US-08-174-185-1
; Sequence 1, Application US/08174185
; Patent No. 5580590
; GENERAL INFORMATION:
; APPLICANT: Hartman, Scott E
; TITLE OF INVENTION: Environmentally Friendly
; TITLE OF INVENTION: Chewing Gum Compositions Containing Elastic Protein-Based
; TITLE OF INVENTION: Polymer
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hill, Steadman & Simpson
; STREET: 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb

; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/174,185
; FILING DATE: 27-DEC-1993
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
US-08-174-185-1

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
| | | | |
DB 1 VPGVG 5

RESULT 4
US-08-212-237-2
; Sequence 2, Application US/08212237
; Patent No. 5606019
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Synthetic Proteins As Implants
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,237
; FILING DATE: 11-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-58847/BIR
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-212-237-2

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
| | | | |

Db 1 VPGVG 5

RESULT 5

US-08-175-155-4
; Sequence 4, Application US/08175155
; Patent No. 5641648
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; APPLICANT: Crissman, John W.
; APPLICANT: Dorman, Mary A.
; TITLE OF INVENTION: Methods for Preparing Synthetic
; TITLE OF INVENTION: Repetitive DNA
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/175,155
; FILING DATE: 29-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I.
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-5/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-175-155-4

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 1 VPGVG 5

RESULT 6

US-08-175-155-9
; Sequence 9, Application US/08175155
; Patent No. 5641648
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; APPLICANT: Crissman, John W.
; APPLICANT: Dorman, Mary A.
; TITLE OF INVENTION: Methods for Preparing Synthetic
; TITLE OF INVENTION: Repetitive DNA
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco

STATE: CA
; COUNTRY: US
; ZIP: 94111
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/175,155
; FILING DATE: 29-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I.
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-5/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-175-155-9

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 1 VPGVG 5

RESULT 7

US-08-468-543-15
; Sequence 15, Application US/08468543
; Patent No. 5726153
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S. et al.
; TITLE OF INVENTION: SYNTHETIC PEPTIDES FOR ARTERIAL IMAGING
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,543
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/398,046
; FILING DATE: 02-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/201,057
; FILING DATE: 24-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,569
; FILING DATE: 16-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694,929

;; FILING DATE: 02-MAY-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/518,215
;; FILING DATE: 03-MAY-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/518,142
;; FILING DATE: 03-MAY-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/189,130
;; FILING DATE: 02-MAY-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Clark, Paul T.
;; REGISTRATION NUMBER: 30,162
;; REFERENCE/DOCKET NUMBER: 04547/002003
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617/542-5070
;; TELEFAX: 617/542-8906
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
US-08-468-543-15

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
Db 1 VPGVG 5

RESULT 8
US-08-477-509B-24
; Sequence 24, Application US/08477509B
; Patent No. 5770697
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A
; APPLICANT: Cappello, Joseph
; APPLICANT: Crissman, John W
; APPLICANT: Dorman, Mary A
; TITLE OF INVENTION: No 5770697el Peptides Comprising Repetitive
; TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,509B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,155
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,049
; FILING DATE: 22-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 06/927,258
;; FILING DATE: 04-NOV-1986
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Trecartin, Richard F.
;; REGISTRATION NUMBER: 31,801
;; REFERENCE/DOCKET NUMBER: A-55186-7/RFT/MTK
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-781-1989
;; TELEFAX: 415-398-3249
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-477-509B-24

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
Db 1 VPGVG 5

RESULT 9
US-08-397-633A-1
; Sequence 1, Application US/08397633A
; Patent No. 5773577
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
; TITLE OF INVENTION: OF ENZYMAIC CROSS-LINKING
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,633A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-397-633A-1

Query Match 100.0%; Score 27; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
| | | | |
Db 1 VPGVG 5

RESULT 10
US-08-397-633A-8
; Sequence 8, Application US/08397633A
; Patent No. 5773577
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
; OF ENZYMIC CROSS-LINKING
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,633A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-397-633A-8

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
| | | | |
Db 1 VPGVG 5

RESULT 11
US-08-707-237A-10
; Sequence 10, Application US/08707237A
; Patent No. 5830713
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; APPLICANT: Crissman, John W.
; APPLICANT: Dorman, Mary A.
; TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC
; REPEPTITIVE DNA
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/707,237A
FILING DATE: 03-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,155
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,049
FILING DATE: 22-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/609,716
FILING DATE: 06-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/269,429
FILING DATE: 09-NOV-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-10/WH
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-707-237A-10

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
| | | | |
Db 1 VPGVG 5

RESULT 12
US-08-483-236-15
; Sequence 15, Application US/08483236
; Patent No. 5939385
; GENERAL INFORMATION:
; APPLICANT: Labroo, Virender
; APPLICANT: Busby, Sharon
; TITLE OF INVENTION: Transglutaminase Cross-Linkable
; POLYPEPTIDES AND METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East

TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-911-364-6

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
| | | | |
Db 1 VPGVG 5

RESULT 15
US-08-735-692-20
Sequence 20, Application US/08735692B
Patent No. 5972406
GENERAL INFORMATION:
APPLICANT: Urry, Dan W.
APPLICANT: Prasad, Karl
TITLE OF INVENTION: Bioplastomers Suitable as Food Product Additives
FILE REFERENCE: BERL-017/01US
CURRENT APPLICATION NUMBER: US/08/735,692B
CURRENT FILING DATE: 1995-10-16
EARLIER APPLICATION NUMBER: 08/423,517
EARLIER FILING DATE: 1995-04-14
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:synthetic
US-08-735-692-20

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
| | | | |
Db 1 VPGVG 5

Search completed: July 25, 2001, 16:49:56
Job time: 67 sec

;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/483,236
;; FILING DATE:
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parker, Gary E
;; REGISTRATION NUMBER: 31-648
;; REFERENCE/DOCKET NUMBER: 93-09c1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6673
;; TELEFAX: 206-442-6678
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-483-236-15

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
Db 1 VPGVG 5

RESULT 13
US-08-469-692-15
; Sequence 15, Application US/08469692
; Patent No. 5955055
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S. et al.
; TITLE OF INVENTION: SYNTHETIC PEPTIDES FOR ARTERIAL IMAGING
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,692
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/398,046
; FILING DATE: 02-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/201,057
; FILING DATE: 24-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,569
; FILING DATE: 16-APR-1993
; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 07/694,929
;; FILING DATE: 02-MAY-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/518,215
;; FILING DATE: 03-MAY-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/518,142
;; FILING DATE: 03-MAY-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/189,130
;; FILING DATE: 02-MAY-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Clark, Paul T.
;; REGISTRATION NUMBER: 30,162
;; REFERENCE/DOCKET NUMBER: 04547/002002
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617/542-5070
;; TELEFAX: 617/542-8906
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; US-08-469-692-15

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
Db 1 VPGVG 5

RESULT 14
US-08-911-364-6
; Sequence 6, Application US/08911364
; Patent No. 5969106
; GENERAL INFORMATION:
; APPLICANT: ROTHSTEIN, Aser
; APPLICANT: KEELY, Fred W.
; TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN
; TITLE OF INVENTION: ELASTIN AND OTHER FIBROUS PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,364
; FILING DATE: 07-AUG-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/023,552
; FILING DATE: 07-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 041082/0104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2001, 16:48:49 ; Search time 22.51 Seconds
(without alignments)
16.920 Million cell updates/sec

Title: US-09-251-638-1

Perfect score: 27

Sequence: 1 VPGVG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_68.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	27	100.0	38	2 T46593	phytoene dehydroge
2	27	100.0	76	2 I45885	elastin - bovine (
3	27	100.0	93	2 C40595	hypothetical prote
4	27	100.0	120	2 A75586	hydrogenase expres
5	27	100.0	127	2 A72712	hypothetical prote
6	27	100.0	139	2 C70680	hypothetical prote
7	27	100.0	162	2 G96671	hypothetical prote
8	27	100.0	170	2 B83664	hypothetical prote
9	27	100.0	181	2 T05925	hypothetical prote
10	27	100.0	195	2 G64359	adenylate kinase (
11	27	100.0	199	2 E69534	DNA polymerase, ba
12	27	100.0	204	2 E72695	adenylate kinase (
13	27	100.0	206	2 D70544	probable amidotran
14	27	100.0	206	2 T45249	probable amidotran
15	27	100.0	213	2 T36699	probable two-compo
16	27	100.0	214	2 T35593	hypothetical prote
17	27	100.0	222	2 J00640	glutamine amidotra
18	27	100.0	231	2 D72532	hypothetical prote
19	27	100.0	246	2 T37169	hypothetical prote
20	27	100.0	255	2 A70778	hypothetical prote
21	27	100.0	267	2 G84319	orotidine-5'-monop
22	27	100.0	271	2 S76871	hypothetical prote
23	27	100.0	274	2 G70898	orotidine-5'-phosp
24	27	100.0	274	2 B49930	orotidine-5'-phosp
25	27	100.0	281	2 G70734	hypothetical prote
26	27	100.0	283	2 B38604	poly(3-hydroxyalka
27	27	100.0	285	2 S29306	poly(3-hydroxyalka
28	27	100.0	293	2 B75441	conserved hypothet
29	27	100.0	294	2 E69968	conserved hypothet

30 27 100.0 301 2 H84192 hypothetical prote
31 27 100.0 303 1 S75782 methanol dehydroge
32 27 100.0 305 2 E75308 amino acid ABC tra
33 27 100.0 305 2 H72289 hypothetical prote
34 27 100.0 306 2 A75316 hypothetical prote
35 27 100.0 318 2 D83725 methanol dehydroge
36 27 100.0 320 1 C69791 methanol dehydroge
37 27 100.0 320 2 G82850 fibrial adhesin p
38 27 100.0 324 2 T18763 hypothetical prote
39 27 100.0 334 2 A31920 collagen sgt-1 pre
40 27 100.0 332 2 T47703 Ca-dependent solut
41 27 100.0 334 2 C70712 hypothetical prote
42 27 100.0 337 2 T34964 probable regulator
43 27 100.0 344 1 WMBE31 38K protein - huma
44 27 100.0 348 2 D84798 probable mitochond
45 27 100.0 358 2 T45934 hypothetical prote

ALIGNMENTS

RESULT 1

T46593

phytoene dehydrogenase [imported] - Mycobacterium marinum (fragment)

C:Species: Mycobacterium marinum

C>Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000

C:Accession: T46593

R:Ramakrishnan, L.; Tran, H.T.; Federspiel, N.A.; Falkow, S.

J. Bacteriol. 179, 5862-5868, 1997

A:Title: A crtB homolog essential for photochromogenicity in Mycobacterium marinum: 1

A:Reference number: 223096; MUID:97440138

A:Accession: T46593

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-38 <RAM>

A:Cross-references: EMBL:U92075; NID:g1928930; PIDN:AAB71427.1; PID:g1928931

A:Experimental source: strain M

C:Genetics:

A:Gene: crtI

Query Match 100.0%; Score 27; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 1 VPGVG 5

RESULT 2

I45885

elastin - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 13-Aug-1999

C:Accession: I45885

R:Rosenbloom, J.

Lab. Invest. 51, 605-623, 1984

A:Title: Biology of disease: Elastin: Relation of protein and gene structure to disea

A:Reference number: I45885; MUID:85059254

A:Accession: I45885

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-76 <ROS>

A:Cross-references: GB:M31891; NID:g163008; PIDN:AAA96416.1; PID:g552319

C:Genetics:

A:Introns: 20/1; 58/1

C:Superfamily: elastin

Query Match 100.0%; Score 27; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 34 VPGVG 38

RESULT 3

C40595

hypothetical protein (mutB 3' region) - Streptomyces cinnamonensis (fragment)

C:Species: Streptomyces cinnamonensis

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Feb-1995

C:Accession: C40595

R:Birch, A.; Leiser, A.; Robinson, J.A.

J. Bacteriol. 175, 3511-3519, 1993

A:Title: Cloning, sequencing, and expression of the gene encoding methylmalonyl-coenzyme

A:Reference number: A40595; MUID:93273720

A:Accession: C40595

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-93 <BIR>

A:Cross-references: GB:L10064

C:Genetics:

A:Start codon: GTG

Query Match

Best Local Similarity 100.0%; Score 27; DB 2; Length 93;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 41 VPGVG 45

RESULT 4

A75586

hydrogenase expression/formation HypA-related protein - Deinococcus radiodurans (strain

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: A75586

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: A75586

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-120 <WHI>

A:Cross-references: GB:AE001863; GB:AE001825; NID:96460670; PIDN:AAF12463.1; PID:9646075

A:Experimental source: strain R1

C:Genetics:

A:Gene: DRA0316

A:Map position: 2

C:Superfamily: hydrogenase accessory protein

Query Match

Best Local Similarity 100.0%; Score 27; DB 2; Length 120;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 67 VPGVG 71

RESULT 5

A72712

hypothetical protein APE112 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000

C:Accession: A72712

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339

A:Accession: A72712

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-127 <RAW>

A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA80097.1; PID:dl043883; PID:g

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE112

C:Superfamily: Aeropyrum pernix hypothetical protein APE112

Query Match

Best Local Similarity 100.0%; Score 27; DB 2; Length 127;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 19 VPGVG 23

RESULT 6

C70680

hypothetical protein RV2437 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: C70680

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: C70680

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-139 <COL>

A:Cross-references: GB:281451; GB:AL123456; NID:g3261662; PIDN:CAB03782.1; PID:e28049

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV2437

Query Match

Best Local Similarity 100.0%; Score 27; DB 2; Length 139;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 132 VPGVG 136

RESULT 7

G96671

hypothetical protein F13011.15 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: G96671

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: G96671

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-162 <STO>

A:Cross-references: GB:AE005173; NID:g5042420; PIDN:AAD38259.1; GSPDB:GN00141

C:Genetics:

A:Gene: F13011.15

A:Map position: 1

Query Match 100.0%; Score 27; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 131 VPGVG 135

RESULT 8

B83664

hypothetical protein BH0114 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000

C:Accession: B83664

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20263314

A:Accession: B83664

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-170 <STO>

A:Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BAB03833.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH0114

Query Match 100.0%; Score 27; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 58 VPGVG 62

RESULT 9

T05925

hypothetical protein - barley

C:Species: Hordeum vulgare (barley)

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999

C:Accession: T05925

R:Hess, W.R.; Golz, R.R.; Boerner, T.

Plant Sci. 133, 191-201, 1998

A:Title: Analysis of randomly selected cDNAs reveals the expression of stress- and defe

A:Reference number: Z15411

A:Accession: T05925

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-181 <HES>

A:Cross-references: EMBL:AJ222779; NID:e1203989; PIDN:CAA10984.1; PID:e1203990

A:Experimental source: cv. Haisa, leaf

Query Match 100.0%; Score 27; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 120 VPGVG 124

RESULT 10

G64359

adenylate kinase (EC 2.7.4.3) MJ0479 [similarity] - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: G64359

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak

rson, J.D.; Sadov, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc

A:Reference number: A64300; MUID:96337999

A:Accession: G64359

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-195 <BUL>

A:Cross-references: GB:U67498; GB:L77117; NID:g1591180; PIDN:AAB98470.1; PID:g1591182

C:Genetics:

A:Map position: FOR421829-422416

A:Start codon: GTG

C:Superfamily: Sulfolobus adenylate kinase

C:Keywords: phosphotransferase

Query Match 100.0%; Score 27; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 14 VPGVG 18

RESULT 11

E69534

DNA polymerase, bacteriophage-type homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 08-Oct-1999

C:Accession: E69534

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artisch, P.; Kaine, B.P.; Sykes,

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch

A:Reference number: A69250; MUID:98049343

A:Accession: E69534

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-199 <KLE>

A:Cross-references: GB:AE000947; GB:AE000782; NID:g2689270; PIDN:AAB88977.1; PID:g264

C:Superfamily: Archaeoglobus probable DNA-polymerase

Query Match 100.0%; Score 27; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 27 VPGVG 31

RESULT 12

E72695

adenylate kinase (EC 2.7.4.3) APE0981 [similarity] - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: E72695
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1
A:Reference number: A72450; MUID:99310339
A:Accession: E72695
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-204 <KAW>
A:Cross-references: DBJ:AP000060; NID:95104188; PIDN:BAA79965.1; PID:95104650
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0981
C:Superfamily: Sulfolobus adenylate kinase
C:Keywords: phosphotransferase

Query Match 100.0%; Score 27; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 15 VPGVG 19

RESULT 13

T36699
Probable amidotransferase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: D70544
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Hooley, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: D70544
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-206 <COL>
A:Cross-references: GB:295586; GB:AL123456; NID:93261785; PIDN:CAB09092.1; PID:g2117235
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: hisH
C:Superfamily: amidotransferase hisH; trpG homology

Query Match 100.0%; Score 27; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 47 VPGVG 51

RESULT 14

T45249
Probable amidotransferase [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 04-Mar-2000
C:Accession: T45249
R:James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, March 1999
A:Reference number: 222949
A:Accession: T45249
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-206 <JAM>

A:Cross-references: EMBL:AL049913; PIDN:CAB43169.1
A:Experimental source: cosmid B1610
C:Genetics:
A:Note: hisH
C:Superfamily: amidotransferase hisH; trpG homology

Query Match 100.0%; Score 27; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 47 VPGVG 51

RESULT 15

T36699
Probable two-component response regulator - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
C:Accession: T36699
R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, April 1999
A:Reference number: Z21597
A:Accession: T36699
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-213 <MUR>
A:Cross-references: EMBL:AL049731; PIDN:CAB41738.1; GSPDB:GN00070; SCOEDB:SCH66.1lc
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCH66.1lc
C:Superfamily: OmpR protein; response regulator homology

Query Match 100.0%; Score 27; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 192 VPGVG 196

Search completed: July 25, 2001, 16:50:26
Job time: 97 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2001, 16:48:49 ; Search time 12.69 Seconds
(without alignments)
13.497 Million cell updates/sec

Title: US-09-251-638-1
Perfect score: 27
Sequence: 1 VPGVG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	27	100.0	93	YMU3_STRCM	Q05072 streptomyce
2	27	100.0	128	COXE_CAEEL	Q20779 caenorhabdi
3	27	100.0	181	Y65L_HORVU	O48609 hordeum vul
4	27	100.0	192	KADA_METIG	P43408 methanococc
5	27	100.0	192	KADA_METJA	P43409 methanococc
6	27	100.0	192	KADA_METTL	P43410 methanococc
7	27	100.0	192	KADA_METVO	P43411 methanococc
8	27	100.0	204	KADA_AERPE	Q9ydd2 aeropyrum p
9	27	100.0	206	HIS5_MYCLE	Q9x7c0 mycobacteri
10	27	100.0	206	HIS5_MYCTU	O06589 mycobacteri
11	27	100.0	211	HIS5_CORGL	O69043 corynebacte
12	27	100.0	222	HIS5_STRCO	P16249 streptomyce
13	27	100.0	255	YM37_MYCTU	Q10519 mycobacteri
14	27	100.0	274	DCOP_MYCTU	P77898 mycobacteri
15	27	100.0	276	DCOP_MYCSM	O08323 mycobacteri
16	27	100.0	281	YNO7_MYCTU	O50658 mycobacteri
17	27	100.0	283	PHAB_PSEOL	P26495 pseudomonas
18	27	100.0	284	STAR_RAT	P97826 rattus norv
19	27	100.0	294	YQXK_BACSU	P38423 bacillus su
20	27	100.0	324	CCSL_CAEEL	P12114 caenorhabdi
21	27	100.0	334	YE96_MYCTU	P71177 mycobacteri
22	27	100.0	374	OMPF_SERMA	O33980 serratia ma
23	27	100.0	393	EX53_MYCTU	Q10899 mycobacteri
24	27	100.0	402	YIN2_STRAM	P32426 streptomyce
25	27	100.0	421	SYH_THETH	P56194 thermus aqu
26	27	100.0	478	GSBH_ARATH	P46416 arabidopsis
27	27	100.0	479	CATA_PSEPU	O59714 pseudomonas
28	27	100.0	482	CATA_ONCVE	Q27710 onchocerca
29	27	100.0	488	U2AF_CAEER	P90727 caenorhabdi
30	27	100.0	496	U2AF_CAEEL	P90978 caenorhabdi
31	27	100.0	510	CP46_RABIT	P14580 oryctolagus
32	27	100.0	511	CP47_RABIT	P14581 oryctolagus
33	27	100.0	519	CBX2_MOUSE	P30658 mus musculu

34 27 100.0 566 1 CAG3_CHICK Q92183 gallus galli
35 27 100.0 582 1 HEMA_MUMPL P19762 mumps virus
36 27 100.0 582 1 HEMA_MUMPM P11235 mumps virus
37 27 100.0 582 1 HEMA_MUMPR P10866 mumps virus
38 27 100.0 583 1 GTB1_MOUSE O08582 mus musculu
39 27 100.0 584 1 GTB1_HUMAN O00178 homo sapien
40 27 100.0 629 1 YS50_MYCTU O05809 mycobacteri
41 27 100.0 649 1 GUND_CLOTUM P04954 clostridium
42 27 100.0 730 1 ELS_HUMAN P15502 homo sapien
43 27 100.0 734 1 VTER_HSVBB P28969 equine herp
44 27 100.0 735 1 VTER_HSV11 P04295 herpes simp
45 27 100.0 747 1 ELS_BOVIN P04985 bos taurus

ALIGNMENTS

RESULT 1
YMU3_STRCM STANDARD; PRT; 93 AA.
AC Q05072;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
*DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHETICAL PROTEIN IN MUTB 3'REGION (ORF-C) (FRAGMENT).
OS Streptomyces cinnamonensis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1900;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3823.5;
RX MEDLINE=93273720; Pubmed=8099072;
RA Birch A., Leiser A., Robinson J.A.;
RT "Cloning, sequencing, and expression of the gene encoding
methyalmalonyl-coenzyme A mutase from Streptomyces cinnamonensis.";
RL J. Bacteriol. 175:3511-3519(1993).
CC -/- SIMILARITY: BELONGS TO THE ARGK FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L10064; AAA03042.1; -;
DR PIR; C40595; C40595.
KW Hypothetical protein.
FT NON_TER 93
SQ SEQUENCE 93 AA; 9858 MW; E667DD645B3845D3 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||

Db 41 VPGVG 45

RESULT 2

COXE_CAEEL STANDARD; PRT; 128 AA.
ID COXE_CAEEL

AC Q20779;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE PROBABLE CYTOCHROME C OXIDASE POLYPEPTIDE VIA PRECURSOR (EC 1.9.3.1).

GN F54D8.2.

OS Caenorhabditis elegans.

CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 CC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Bentley D.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 CC 4 FERRICYTOCHROME C.
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIA FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U12966; AAA20614.1; -
 DR WormPep; F54D8.2; CE01308.
 DR InterPro; IPR001349; -
 DR Pfam; PF02046; COX6A; 1.
 DR PROSITE; PS01329; COX6A; 1.
 KW Oxidoreductase; Inner membrane; Mitochondrion; Transit peptide.
 FT TRANSIT 1 128 MITOCHONDRION (POTENTIAL).
 FT CHAIN ? 128 PROBABLE CYTOCHROME C OXIDASE POLYPEPTIDE
 FT VIA.
 SQ SEQUENCE 128 AA; 14743 MW; EF4EA56A1CE6A233 CRC64;
 Query Match 100.0%; Score 27; DB 1; Length 128;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPGVG 5
 Db 116 VPGVG 120
 RESULT 3
 ID Y65L_HORVU STANDARD; PRT; 181 AA.
 AC O48609;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 DE YCF65-LIKE PROTEIN PRECURSOR.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae;
 OC Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. HAISA; TISSUP-Leaf;
 RA Hess W.R., Golz R., Boerner T.;
 RT "Analysis of randomly selected cDNAs reveals the expression of stress-
 RT and defence-related genes in the barley mutant albobastrians.";
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE YCF65 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ222779; CAA10984.1; -
 KW Hypothetical protein; Chloroplast; Transit peptide.
 FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
 FT CHAIN ? 181 YCF65-LIKE PROTEIN.
 SQ SEQUENCE 181 AA; 19865 MW; B02DAC3792F7E885 CRC64;
 Query Match 100.0%; Score 27; DB 1; Length 181;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPGVG 5
 Db 120 VPGVG 124
 RESULT 4
 ID KADA_METIG STANDARD; PRT; 192 AA.
 AC P43408;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
 GS ADKA OR ADK.
 OS Methanococcus igneus.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus
 OX NCBI_TaxID=2189;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97208879; PubMed-9055821;
 RA Ferber D.M., Haney P.J., Berk H., Lynn D., Konisky J.;
 RT "The adenylate kinase genes of M. voltae, M. thermolithotrophicus, M.
 RT Jannaschii, and M. igneus define a new family of adenylate kinases.";
 RL Gene 185:239-244(1997).
 RN [2]
 RP SEQUENCE OF 1-30.
 RX MEDLINE-95286473; PubMed-7768791;
 RA Rusnak P., Haney P., Konisky J.;
 RT "The adenylate kinases from a mesophilic and three thermophilic
 RT methanogenic members of the Archaea.";
 RL J. Bacteriol. 177:2977-2981(1995).
 CC -!- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
 CC -!- SUBUNIT: MONOMER (PROBABLE).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- MISCELLANEOUS: ACTIVE IN THE TEMPERATURE RANGE OF 70 TO 90 DEGREES
 CC CELSIUS.
 CC -!- SIMILARITY: BELONGS TO THE ARCHAEL ADENYLATE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U39881; AAC44862.1; -
 KW Transferase; Kinase; ATP-binding.
 FT NP_BIND 10 18 ATP (BY SIMILARITY).
 FT CONFLICT 8 8 V -> I (IN REF. 2).
 FT CONFLICT 20 20 T -> L (IN REF. 2).
 SQ SEQUENCE 192 AA; 21400 MW; AB2EC9C9DB905E75 CRC64;
 Query Match 100.0%; Score 27; DB 1; Length 192;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 DB 11 VPGVG 15

RESULT 5

KADA_METJA STANDARD; PRT; 192 AA.
 AC P43409;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
 GN ADKA OR ADK OR MJ0479.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97208879; PubMed=9055821;
 RA Ferber D.M., Haney P.J., Berk H., Lynn D., Konisky J.;
 RT "The adenylate kinase genes of M. voltae, M. thermolithotrophicus, M.
 jannaschii, and M. igneus define a new family of adenylate kinases.";
 RL Gene 185:239-244(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Sutt C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Bult C.J., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Herrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 RN [3]
 RP SEQUENCE OF 1-30.
 RX MEDLINE=95286473; PubMed=7768791;
 RA Rusnak P., Haney P., Konisky J.;
 RT "The adenylate kinases from a mesophilic and three thermophilic
 methanogenic members of the Archaea.";
 RL J. Bacteriol. 177:2977-2981(1995).
 CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
 CC -1- SUBUNIT: MONOMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- MISCELLANEOUS: ACTIVE IN THE TEMPERATURE RANGE OF 70 TO 90 DEGREES
 CELSIUS.
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEL ADENYLATE KINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 DR EMBL; U39882; AAC44863.1;
 DR EMBL; U67498; AAB98470.1; ALT_INIT.
 DR TIGR; MJ0479;
 KW Transferase; Kinase; ATP-binding.
 FT NP_BIND 10 18 ATP (BY SIMILARITY).
 FT CONFLICT 16 16 S->G (IN REF. 3).
 SQ SEQUENCE 192 AA; 21772 MW; ECD533AD4C85D99E CRC64;
 Query Match 100.0%; Score 27; DB 1; Length 192;

QY 1 VPGVG 5
 DB 11 VPGVG 15

Best Local Similarity 100.0%; Pred. No. le+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 DB 11 VPGVG 15

RESULT 6

KADA_METTL STANDARD; PRT; 192 AA.
 AC P43410;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
 GN ADKA OR ADK.
 OS Methanococcus thermolithotrophicus.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2186;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97208879; PubMed=9055821;
 RA Ferber D.M., Haney P.J., Berk H., Lynn D., Konisky J.;
 RT "The adenylate kinase genes of M. voltae, M. thermolithotrophicus, M.
 jannaschii, and M. igneus define a new family of adenylate kinases.";
 RL Gene 185:239-244(1997).
 RN [2]
 RP SEQUENCE OF 1-20.
 RX MEDLINE=95286473; PubMed=7768791;
 RA Rusnak P., Haney P., Konisky J.;
 RT "The adenylate kinases from a mesophilic and three thermophilic
 methanogenic members of the Archaea.";
 RL J. Bacteriol. 177:2977-2981(1995).
 CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
 CC -1- SUBUNIT: MONOMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- MISCELLANEOUS: ACTIVE IN THE TEMPERATURE RANGE OF 60 TO 80 DEGREES
 CELSIUS.
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEL ADENYLATE KINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 DR EMBL; U39880; AAC44864.1;
 DR Transferase; Kinase; ATP-binding.
 FT NP_BIND 10 18 ATP (BY SIMILARITY).
 SQ SEQUENCE 192 AA; 21461 MW; 72233378B43320B1 CRC64;
 Query Match 100.0%; Score 27; DB 1; Length 192;
 Best Local Similarity 100.0%; Pred. No. le+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 DB 11 VPGVG 15

RESULT 7

KADA_METVO STANDARD; PRT; 192 AA.
 AC P43411;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
 Query Match 100.0%; Score 27; DB 1; Length 192;

GN ADKA OR ADK.
 OS Methanococcus voltae.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2188;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PS;
 RX MEDLINE=97208879; PubMed=9055821;
 RA Ferber D.M., Haney P.J., Berk H., Lynn D., Konisky J.;
 RT "The adenylate kinase genes of M. voltae, M. thermophilic and three thermophilic
 RT methanogenic members of the Archaea.";
 RL J. Bacteriol. 177:2977-2981(1995).
 RL Janaschil, and M. igneus define a new family of adenylate kinases.";
 RN Gene 185:239-244(1997).
 RP [2]
 RP SEQUENCE OF 1-38.
 RX MEDLINE=95286473; PubMed=7768791;
 RA Rusnak P., Haney P., Konisky J.;
 RT "The adenylate kinases from a mesophilic and three thermophilic
 RT methanogenic members of the Archaea.";
 RL J. Bacteriol. 177:2977-2981(1995).
 CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
 CC -1- SUBUNIT: MONOMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- MISCELLANEOUS: ACTIVE IN THE TEMPERATURE RANGE OF 30 TO 40 DEGREES
 CC CELSIUS.
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEAL ADENYLATE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U39879; AAC4865.1;
 KW Transferase; Kinase; ATP-binding.
 FT NP_BIND 10 18 ATP (BY SIMILARITY).
 FT CONFLICT 31 31 G -> GG (IN REF. 2).
 SQ SEQUENCE 192 AA; 21303 MW; 952ABCD1788D6A8E CRC64;

Query Match 100.0%; Score 27; DB 1; Length 192;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 Db 11 VPGVG 15

RESULT 8 KADA_AERPE

ID KADA_AERPE STANDARD; PRT; 204 AA.
 AC Q9YDD2;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
 GN ADKA OR APE0981.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
 OC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,

RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEAL ADENYLATE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AP000060; BAA79965.1;
 KW Transferase; Kinase; ATP-binding.
 FT NP_BIND 14 22 ATP (BY SIMILARITY).
 SQ SEQUENCE 204 AA; 22175 MW; 279EE5F8BDB81D13 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 204;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 Db 15 VPGVG 19

RESULT 9 HIS5_MYCLE

ID HIS5_MYCLE STANDARD; PRT; 206 AA.
 AC Q9X7C0;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE AMIDOTRANSFERASE HISH (EC 2.4.2.-).
 GN HISH OR MLCB1610.23.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seeger K.J., Harris D., James K.D., Parkhill J., Barrell B.G.,
 RA Rajandream M.A.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES AN AMIDOTRANSFERASE REACTION THAT GENERATES
 CC MIDAZOLE-GLYCEROL PHOSPHATE AND 5-AMINOIMIDAZOL-4-CARBOXAMIDE
 CC RIBONUCLEOTIDE, WHICH IS USED FOR PURINE SYNTHESIS.
 CC -1- PATHWAY: FIFTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE HISH FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).

EMBL; AL049913; CAB43169.1;
 InterPro: IPR000991;
 DR Pfam; PF00117; CATase; 1.

DR PROSITE; PS00442; CATASE_TYPE_I; 1.
 KW Histidine biosynthesis; Transferase; Glutamine amidotransferase.
 FT ACT_SITE 83 83 BY SIMILARITY.
 FT ACT_SITE 187 187 BY SIMILARITY.
 FT ACT_SITE 189 189 BY SIMILARITY.

SQ SEQUENCE 206 AA; 21652 MW; 72D6994084F81536 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 47 VPGVG 51

RESULT 10

HIS5_MYCTU STANDARD; PRT; 206 AA.
AC 006589;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE AMIDOTRANSFERASE HISH (EC 2.4.2.-).
GN HISH OR RV1602 OR MTCY336.02C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
CC -1- FUNCTION: CATALYZES AN AMIDOTRANSFERASE REACTION THAT GENERATES
CC IMIDAZOLE-GLYCEROL PHOSPHATE AND 5-AMINOIMIDAZOL-4-CARBOXAMIDE
CC RIBONUCLEOTIDE, WHICH IS USED FOR PURINE SYNTHESIS.
CC -1- PATHWAY: FIFTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE HISH FAMILY.
CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----

EMBL: Z95586; CAB09092.1; -
DR TubercuList; RV1602; -
DR InterPro: IPR000991; -
DR Pfam: PF00117; GATase; 1.
DR PROSITE; PS00442; GATASE_TYPE_I; 1.
KW Histidine biosynthesis; Transferase; Glutamine amidotransferase.
FT ACT_SITE 83 83 BY SIMILARITY.
FT ACT_SITE 187 187 BY SIMILARITY.
FT ACT_SITE 189 189 BY SIMILARITY.
SQ SEQUENCE 206 AA; 21418 MW; 4B2AF9C61BEE1447 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

DB 47 VPGVG 51
|||||

RESULT 11

HIS5_CORGL STANDARD; PRT; 211 AA.
AC 069043;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE AMIDOTRANSFERASE HISH (EC 2.4.2.-).
GN HISH.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13059 / AS019;
RA Juns S.I., Han M.S., Park Y.J., Lee S.K., Lee M.S.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES AN AMIDOTRANSFERASE REACTION THAT GENERATES
CC IMIDAZOLE-GLYCEROL PHOSPHATE AND 5-AMINOIMIDAZOL-4-CARBOXAMIDE
CC -1- PATHWAY: FIFTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE HISH FAMILY.
CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----

EMBL: AF060558; AAC15231.1; -
DR InterPro: IPR000991; -
DR Pfam: PF00117; GATase; 1.
DR PROSITE; PS00442; GATASE_TYPE_I; 1.
KW Histidine biosynthesis; Transferase; Glutamine amidotransferase.
FT ACT_SITE 82 82 BY SIMILARITY.
FT ACT_SITE 190 190 BY SIMILARITY.
FT ACT_SITE 192 192 BY SIMILARITY.
SQ SEQUENCE 211 AA; 23139 MW; A935FEC4C2B48AB1 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 47 VPGVG 51

RESULT 12

HIS5_STRCO STANDARD; PRT; 222 AA.
AC P16249;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE AMIDOTRANSFERASE HISH (EC 2.4.2.-).
GN HISH OR SC4G6.20C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]

Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
Nature 393:537-544(1998).

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Z70692; CAA94667.1; -
DR TuberculList; RV2237; -
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 222 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
SQ SEQUENCE 255 AA; 29067 MW; B0F422277C6B46A6 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 14 VPGVG 18

RESULT 14
DCOP_MYCTU
ID DCOP_MYCTU STANDARD; PRT; 274 AA.
AC P77898; P42610;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE OROTIDINE 5'-PHOSPHATE DECARBOXYLASE (EC 4.1.1.23) (OMP
DE DECARBOXYLASE) (OMPDCASE).
GN PYRF OR URAA OR RV1385 OR MTCY21B4.02.
OS Mycobacterium tuberculosis, and Mycobacterium bovis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Radoock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies K., Devlin K., Feitwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsbay T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
Nature 393:537-544(1998).
[2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.Dovis; STRAIN=BCG;
RX MEDLINE=94042902; PubMed=8226675;
RA Aldovini A., Hussen R.N., Young R.A.;
"The uraA locus and homologous recombination in Mycobacterium bovis
BCG".
RL J. Bacteriol. 175:7282-7289(1993).
CC -1- CATALYTIC ACTIVITY: OROTIDINE-5'-PHOSPHATE = UMP + CO(2).
CC -1- PATHWAY: SIXTH AND LAST STEP IN THE BIOSYNTHESIS OF PYRIMIDINES.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; 280108; CAB02190.1; -;
 DR EMBL; 001072; AAC43182.1; -;
 DR TuberculList; RV1385; -;
 DR InterPro; IPR001754; -;
 DR Pfam; PF00215; OMPdecase; 1;
 DR PROSITE; PS00156; OMPDECASE; 1;
 DR Pyrimidine biosynthesis; Lyase; Decarboxylase.
 KW ACT_SITE 95 BY SIMILARITY.
 FT ACT_SITE 95
 SQ SEQUENCE 274 AA; 27377 MW; 369BAE076FB3D143 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 274;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 |||||
 DB 215 VPGVG 219

RESULT 15
 DCOP_MYCSM
 ID DCOP_MYCSM STANDARD; PRT; 276 AA.
 AC O08323;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE OROTIDINE 5'-PHOSPHATE DECARBOXYLASE (EC 4.1.1.23) (OMP
 DE DECARBOXYLASE) (OMPDCASE).
 GN PYRF.
 OS Mycobacterium smegmatis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1772;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 700084 / MC(2)155;
 RX MEDLINE=97312763; PubMed=9169204;
 RA Knipfer N., Seth A., Shrader T.E.;
 RT "Unmarked gene integration into the chromosome of Mycobacterium
 RL smegmatis via precise replacement of the pyrF gene.";
 RL Plasmid 37:129-140(1997).
 CC -1- CATALYTIC ACTIVITY: OROTIDINE-5'-PHOSPHATE = UMP + CO(2).
 CC -1- PATHWAY: SIXTH AND LAST STEP IN THE BIOSYNTHESIS OF PYRIMIDINES.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U91572; AAB50157.1; -;
 DR InterPro; IPR001754; -;
 DR Pfam; PF00215; OMPdecase; 1;
 DR PROSITE; PS00156; OMPDECASE; 1;
 DR Pyrimidine biosynthesis; Lyase; Decarboxylase.
 KW ACT_SITE 95 BY SIMILARITY.
 FT ACT_SITE 95
 SQ SEQUENCE 276 AA; 27815 MW; 41437843A3E2A896 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 276;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 |||||
 DB 221 VPGVG 225

Search completed: July 25, 2001, 16:51:25
 Job time: 156 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2001, 16:48:49 ; Search time 33.32 Seconds
(without alignments)
19.854 Million cell updates/sec

Title: US-09-251-638-1

Perfect score: 27

Sequence: 1 VPGVG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_16:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_unclassified:*

13: sp_vertebrate:*

14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	27	100.0	38	005422	005422 mycobacteri
2	27	100.0	76	028100	028100 bos taurus
3	27	100.0	100	09H402	09H402 homo sapien
4	27	100.0	120	09RYJ6	09RYJ6 delnococtus
5	27	100.0	127	09XD00	09XD00 aeropyrum p
6	27	100.0	130	018832	018832 equus cabal
7	27	100.0	138	09HHZ4	09HHZ4 halobacteri
8	27	100.0	139	09H12	09H12 mycobacteri
9	27	100.0	141	066301	066301 unidentifie
10	27	100.0	141	066330	066330 unidentifie
11	27	100.0	141	066336	066336 unidentifie
12	27	100.0	141	066364	066364 unidentifie
13	27	100.0	141	066365	066365 unidentifie
14	27	100.0	141	066367	066367 unidentifie
15	27	100.0	143	09N2X5	09N2X5 caenorhabdi
16	27	100.0	162	09XIR0	09XIR0 arabidopsis
17	27	100.0	167	09V506	09V506 drosophila
18	27	100.0	169	09PVQ0	09PVQ0 xenopus lae
19	27	100.0	170	09KGF1	09KGF1 bacillus ha

20	27	100.0	176	11	09JUJ3	Q9JJU3 mus musculu
21	27	100.0	181	4	09UJK2	Q9UJK2 homo sapien
22	27	100.0	183	10	09XGW3	Q9XGW3 oryza sativ
23	27	100.0	189	3	09US43	Q9US43 schizosach
24	27	100.0	199	1	028007	028007 archaesoglob
25	27	100.0	202	1	09HK44	09HK44 thermoplasma
26	27	100.0	213	2	09X943	09X943 streptomyce
27	27	100.0	214	5	045681	045681 caenorhabdi
28	27	100.0	223	9	09T111	09T111 lactobacill
29	27	100.0	228	4	09NV42	09NV42 homo sapien
30	27	100.0	231	1	09Y906	09Y906 aeropyrum p
31	27	100.0	234	7	031379	031379 cyprinus ca
32	27	100.0	234	7	031380	031380 cyprinus ca
33	27	100.0	237	2	09RDA0	09RDA0 streptomyce
34	27	100.0	237	5	09VCU6	09VCU6 drosophila
35	27	100.0	246	2	09SIN7	09SIN7 streptomyce
36	27	100.0	250	4	09H5H1	09H5H1 homo sapien
37	27	100.0	258	4	09UMF5	09UMF5 homo sapien
38	27	100.0	265	2	09KI91	09KI91 bacillus an
39	27	100.0	267	1	09P9M3	09P9M3 halobacteri
40	27	100.0	267	5	09VHH4	09VHH4 drosophila
41	27	100.0	271	2	074665	074665 synecocyst
42	27	100.0	273	2	09K188	09K188 bacillus ce
43	27	100.0	283	2	09R9W3	09R9W3 pseudomonas
44	27	100.0	285	2	09Z3Y0	09Z3Y0 pseudomonas
45	27	100.0	285	2	09X5X8	09X5X8 pseudomonas

ALIGNMENTS

RESULT 1

ID 005422 PRELIMINARY; PRT; 38 AA.

AC 005422;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

DE PHYTOENE DEHYDROGENASE (FRAGMENT).

GN CRTI.

OS Mycobacterium marinum.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1781;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=M;

RX MEDLINE=97440138; PubMed=9294446;

RA Ramakrishnan L., Tran H.T., Federspiel N.A., Falkow S.;

RT "A crtB homolog essential for photochromogenicity in Mycobacterium

RT marinum: isolation, characterization, and gene disruption via

RT homologous recombination.";

RL J. Bacteriol. 179:5862-5868(1997).

DR EMBL; U92075; AAB71427.1; -

FT NON_TER 1 1

SQ SEQUENCE 38 AA; 3986 MW; 6E46332707CCDCAC CRC64;

Query Match 100.0%; Score 27; DB 2; Length 38;

Best Local Similarity 100.0%; Pred. No. 91;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

|||||

Db 1 VPGVG 5

RESULT 2

ID 028100 PRELIMINARY; PRT; 76 AA.

AC 028100;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ELASTIN (FRAGMENT).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85059254; PubMed=6150137;
 RT "Elastin: relation of protein and gene structure to disease.";
 RL Lab. Invest. 51:605-623(1984).
 DR EMBL; M31891; AAA96416.1;
 DR EMBL; M31893; AAA96416.1; JOINED.
 DR EMBL; M31892; AAA96416.1; JOINED.
 FT NON_TER 1
 SQ SEQUENCE 76 AA; 6619 MW; B683379DAE87B202 CRC64;

Query Match 100.0%; Score 27; DB 6; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 Db 34 VPGVG 38

RESULT 3
 Q9H402
 ID Q9H402 PRELIMINARY; PRT; 100 AA.
 AC Q9H402;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE DJ583P15.9 (KIAA1088) (FRAGMENT).
 GN BK318A7.3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matthews L.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL121845; CAC16546.1;
 FT NON_TER 1
 SQ SEQUENCE 100 AA; 10169 MW; 44BA216B2A028007 CRC64;

Query Match 100.0%; Score 27; DB 4; Length 100;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 Db 15 VPGVG 19

RESULT 4
 Q9RYJ6
 ID Q9RYJ6 PRELIMINARY; PRT; 120 AA.
 AC Q9RYJ6;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE HYDROGENASE EXPRESSION/FORMATION HYPA-RELATED PROTEIN.
 GN DRA0316.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=RL;
 RX MEDLINE=20036896; PubMed=10567466;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans RL.";
 RL Science 286:1571-1577(1999).
 DR EMBL; AE001863; AAF12463.1;
 DR TIGR; DRA0316;
 DR InterPro; IPR000688;
 DR Pfam; PF01155; HypA; 1.
 SQ SEQUENCE 120 AA; 12680 MW; F4D6904A721157D4 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 120;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 Db 67 VPGVG 71

RESULT 5
 Q9YD00
 ID Q9YD00 PRELIMINARY; PRT; 127 AA.
 AC Q9YD00;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 13.6 KDA PROTEIN APE1112.
 GN APE1112.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
 OC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL; AP000060; BAA80097.1;
 KW Hypothetical protein.
 SQ SEQUENCE 127 AA; 13575 MW; F4900D42B1099BD7 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 127;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 Db 19 VPGVG 23

RESULT 6
 O18832
 ID O18832 PRELIMINARY; PRT; 130 AA.
 AC O18832;

DT 01-NOV-1998 (TEMBLrel. 08, Created)
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
 DE 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
 DE AGGRECAN CORE PROTEIN (CARTILAGE-SPECIFIC PROTEOGLYCAN CORE PROTEIN)
 DE (CSPCP) (FRAGMENT).
 GN AGC1.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OC NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-CHONDROCYTES;
 RX MEDLINE=98209637; PubMed=9550267;
 RA Flannery C.R., Little C.B., Caterson B.;
 RA "Molecular cloning and sequence analysis of the aggrecan interglobular
 RT domain from porcine, equine, bovine and ovine cartilage: comparison of
 RT proteinase-susceptible regions and sites of keratan sulfate
 RT substitution.";
 RL Matrix Biol. 16:507-511(1997).
 CC -1- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
 CC MATRIX OF CARTILAGINOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
 CC IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
 CC HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A
 CC REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.
 CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
 CC SIMILARITY).
 CC -1- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
 CC TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,
 CC MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
 CC CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS
 CC THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
 CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
 CC AND G3.
 CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
 CC CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
 DR EMBL; AF019756; AAC48798.1; -;
 DR InterPro; IPR000538; -;
 DR PROSITE; PS01241; LINK; PARTIAL.
 KW Glycoprotein; Cartilage; Proteoglycan; Repeat.
 FT NON_TER 1
 FT DOMAIN <1 10 LINK 2.
 FT DOMAIN <1 10 G1-B'.
 FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 130 130
 SQ SEQUENCE 130 AA; 13830 MW; 082F79B5AE78B53E CRC64;

 Query Match 100.0%; Score 27; DB 6; Length 130;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0;
 QY 1 VPGVG 5
 DB 82 VPGVG 86

 RESULT 7
 Q9HHZ4 PRELIMINARY; PRT; 138 AA.
 AC Q9HHZ4;
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
 DE VNG6157H.
 GN VNG6157H.
 OS Halobacterium sp. (strain NRC-1).
 OG Plasmid pNRC200.
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
 OC Halobacterium.
 OC NCBI_TaxID=64091;

RN SEQUENCE FROM N.A.
 RP MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithauser B., Kellier K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Fohlschroder M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
 RT "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL; AE005152; AAG20828.1; -;
 KW Plasmid.
 SQ SEQUENCE 138 AA; 14950 MW; 3E347A1BB44F2DAC CRC64;

 Query Match 100.0%; Score 27; DB 1; Length 138;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPGVG 5
 DB 101 VPGVG 105

 RESULT 8
 P71912 PRELIMINARY; PRT; 139 AA.
 ID P71912;
 AC P71912;
 DT 01-FEB-1997 (TEMBLrel. 02, Created)
 DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
 DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
 DE HYPOTHETICAL 15.3 KDA PROTEIN.
 GN RV2437 OR MTCY428.09C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 DR EMBL; Z81451; CAB03782.1; -;
 DR TubercuList; RV2437; -;
 KW Hypothetical protein.
 SQ SEQUENCE 139 AA; 15315 MW; 842BF116C0E102EC CRC64;

 Query Match 100.0%; Score 27; DB 2; Length 139;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPGVG 5
 DB 132 VPGVG 136

 RESULT 9
 O66301 PRELIMINARY; PRT; 141 AA.
 ID O66301

O66301;
 AC 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH.
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 OX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ohkuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB011855; BAA28392.1; -
 DR HSSP; P00456; ICP2.
 DR InterPro: IPR000392; -
 DR Pfam; PF00142; fer4_NifH; 1.
 DR PRINTS; PR00091; NITROGNASEII.
 DR PROSITE; PS00746; NIFH_FRXC_1; 1.
 DR PROSITE; PS00692; NIFH_FRXC_2; 1.
 FT NON_TER 1
 FT NON_TER 141
 SQ SEQUENCE 141 AA; 14882 MW; 02454ECD55EDBF40 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 |
 |
 |
 |
 Db 74 VPGVG 78

RESULT 10
 O66330
 ID O66330 PRELIMINARY; PRT; 141 AA.
 AC O66330;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH.
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 OX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ohkuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB011900; BAA28435.1; -
 DR HSSP; P00456; ICP2.
 DR InterPro: IPR000392; -
 DR Pfam; PF00142; fer4_NifH; 1.
 DR PRINTS; PR00091; NITROGNASEII.
 DR PROSITE; PS00746; NIFH_FRXC_1; 1.
 DR PROSITE; PS00692; NIFH_FRXC_2; 1.
 FT NON_TER 1
 FT NON_TER 141
 SQ SEQUENCE 141 AA; 14632 MW; 7251716CF85C6752 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 |
 |
 |
 |
 Db 74 VPGVG 78

RESULT 11
 O66336

ID O66336 PRELIMINARY; PRT; 141 AA.
 AC O66336;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH.
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 OX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ohkuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB011906; BAA28441.1; -
 DR HSSP; P00456; ICP2.
 DR InterPro: IPR000392; -
 DR Pfam; PF00142; fer4_NifH; 1.
 DR PRINTS; PR00091; NITROGNASEII.
 DR PROSITE; PS00746; NIFH_FRXC_1; 1.
 DR PROSITE; PS00692; NIFH_FRXC_2; 1.
 FT NON_TER 1
 FT NON_TER 141
 SQ SEQUENCE 141 AA; 14910 MW; D2E06148699D28C3 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 |
 |
 |
 |
 Db 74 VPGVG 78

RESULT 12
 O66364
 ID O66364 PRELIMINARY; PRT; 141 AA.
 AC O66364;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH.
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 OX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ohkuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB011954; BAA28487.1; -
 DR HSSP; P00456; ICP2.
 DR InterPro: IPR000392; -
 DR Pfam; PF00142; fer4_NifH; 1.
 DR PROSITE; PS00746; NIFH_FRXC_1; 1.
 DR PROSITE; PS00692; NIFH_FRXC_2; 1.
 FT NON_TER 1
 FT NON_TER 141
 SQ SEQUENCE 141 AA; 14780 MW; E069FA235F62484B CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 |
 |
 |
 |
 Db 74 VPGVG 78

RESULT 13
 O66365

ID O66365 PRELIMINARY; PRT; 141 AA.
 AC O66365;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH.
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 OX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ohkuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB011955; BAA28488.1; -
 DR HSSP; P00456; ICP2;
 DR InterPro; IPR000392; -
 DR Pfam; PF00142; fer4_NifH; 1.
 DR PROSITE; PS00746; NIFH_FRXC_1; 1.
 DR PROSITE; PS00692; NIFH_FRXC_2; 1.
 FT NON_TER 1-1
 FT NON_TER 141
 SQ SEQUENCE 141 AA; 14786 MW; 2BC10A2BF5C05E2E CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 Db 74 VPGVG 78

RESULT 14

O66367 PRELIMINARY; PRT; 141 AA.
 ID O66367;
 AC O66367;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH.
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 OX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ohkuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB011957; BAA28490.1; -
 DR HSSP; P00456; ICP2;
 DR InterPro; IPR000392; -
 DR Pfam; PF00142; fer4_NifH; 1.
 DR PROSITE; PS00746; NIFH_FRXC_1; 1.
 DR PROSITE; PS00692; NIFH_FRXC_2; 1.
 FT NON_TER 1-1
 FT NON_TER 141
 SQ SEQUENCE 141 AA; 14505 MW; 7A2B008BE564C29C CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 Db 74 VPGVG 78

RESULT 15

O66367 PRELIMINARY; PRT; 143 AA.
 ID O66367;
 AC O66367;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH.
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 OX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ohkuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB011957; BAA28490.1; -
 DR HSSP; P00456; ICP2;
 DR InterPro; IPR000392; -
 DR Pfam; PF00142; fer4_NifH; 1.
 DR PROSITE; PS00746; NIFH_FRXC_1; 1.
 DR PROSITE; PS00692; NIFH_FRXC_2; 1.
 FT NON_TER 1-1
 FT NON_TER 141
 SQ SEQUENCE 141 AA; 14505 MW; 7A2B008BE564C29C CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 Db 74 VPGVG 78

AC Q9N2X5;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Y92C3B.B PROTEIN.
 GN Y92C3B.B.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; Pubmed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC024875; AAF60885.1; -
 DR InterPro; IPR000504; -
 DR Pfam; PF00076; rim; 1.
 SQ SEQUENCE 143 AA; 16000 MW; CDFD9C94F94A00D7 CRC64;

Query Match 100.0%; Score 27; DB 5; Length 143;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGVG 5
 Db 95 VPGVG 99

Search completed: July 25, 2001, 16:51:06
 Job time: 137 sec

